SOME RESULTS ON SUMS AND PRODUCTS

A Thesis
Presented to
The Academic Faculty

by

Christopher Ian Pryby

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
December 2014

SOME RESULTS ON SUMS AND PRODUCTS

Approved by:

Professor Ernie Croot, Advisor
School of Mathematics
Georgia Institute of Technology

Professor Michael Lacey
School of Mathematics
Georgia Institute of Technology

Professor Neil Lyall
Department of Mathematics
University of Georgia

Professor William T. Trotter
School of Mathematics
Georgia Institute of Technology

Professor Xingxing Yu
School of Mathematics
Georgia Institute of Technology

Date Approved: 13 November 2014

To Brittany,

Dalí,

and Dedekind

$$2 + 2 = 2 \times 2 = 2^2$$

ACKNOWLEDGEMENTS

I would like to thank my advisor, Prof. Ernie Croot, for his patience and his encouragement in my progress in my graduate student career. His advice and insight have aided me countless times in my path at Georgia Tech.

I would like to thank my committee members for their advice and feedback on the present work. Additionally, Profs. Neil Lyall and Michael Lacey were instrumental in training me in analysis during my undergraduate and graduate studies at the University of Georgia and Georgia Tech.

I would also like to thank my fellow graduate students and co-authors Albert Bush and Gagik Amirkhanyan for their contributions to the present work. I appreciate the many helpful and engaging discussions we have had.

I would like to thank my cohort of budding mathematicians during my undergraduate studies at the University of Georgia, especially Meredith Casey and Alex Rice. We had great times in Prof. Ted Shifrin's multivariable mathematics class, which started us all down paths toward mathematical careers. Thank you also to Prof. Shifrin himself, who has crafted a brilliant introduction to higher mathematics in the University of Georgia math program. This program is why I became a mathematician: my interest in deeper math was piqued from the first day of my calculus with theory course, and it has continued onward ever since.

I would like to thank Anthony Coulter, my friend since high school and a fellow lover of mathematics, for his support, encouragement, and our fascinating conversations. But not for his best man speech.

Finally I would like to thank my loving and supportive family, especially my parents, grandparents, and my wife Brittany.

TABLE OF CONTENTS

SUMMARY

We demonstrate new results in additive combinatorics, including a proof of a conjecture by J. Solymosi: for every $\varepsilon > 0$, there exists $\delta > 0$ such that, given n^2 points in a grid formation in $\mathbb{R}^2$, if L is a set of lines in general position such that each line intersects at least $n^{1-\delta}$ points of the grid, then $|L| < n^{\varepsilon}$. This result implies a conjecture of Gy. Elekes regarding a uniform statistical version of Freiman's theorem for linear functions with small image sets.

CHAPTER I

INTRODUCTION

1.1 The Arithmetic of Sets

Let G be a group with binary operation $+$, and let A, B be subsets of G. We define the *sum set* of A and B by

$$A + B := \{a + b : a \in A, b \in B\}.$$

We can further define the *iterated sum set* by $1A := A$ and, for $k > 1$,

$$kA := (k-1)A + A.$$

The related notion of the *difference set* of A and B can also be defined by

$$A - B := \{a - b : a \in A, b \in B\}.$$

If A and B are subsets of a ring R, then we can define their sum set as above using the addition operation on R. We can also define the *product set* of A and B by

$$A.B := \{a \cdot b : a \in A, b \in B\},$$

where $\cdot$ is the multiplication operation on R. If B is a subset of the group of units of R, then we can define the *ratio set* of A and B by

$$A/B := \left\{a \cdot b^{-1} : a \in A, b \in B\right\}.$$

The *iterated product set* is defined by $A^{(1)} = A$ and

$$A^{(k)} := A^{(k-1)}.A$$

for $k > 1$.

Two major objects of study in the field of additive combinatorics are the ratios $\frac{|A+A|}{|A|}$ and $\frac{|A.A|}{|A|}$, which give measures of the *growth* of a finite set A under the operation of taking its sum set with itself and its product set with itself. We may also study the related ratios $\frac{|A-A|}{|A|}$ and $\frac{|A/A|}{|A|}$. If A is subset of an additive group, [72] denotes the ratio $\frac{|A+A|}{|A|}$ by $\sigma[A]$, the *doubling constant*, and the ratio $\frac{|A-A|}{|A|}$ by $\delta[A]$, the *difference constant*. (In principle, the same symbols can be used if the operation of G is written multiplicatively.)

It is not hard to see that the minimum possible doubling constant of a set A is 1: this occurs precisely when A is a coset of a finite subgroup of the ambient group G [72]. On the other hand, the maximum possible doubling constant is $|A|$: this would occur if $a + b \neq c + d$ whenever $(a, b) \neq (c, d)$. (In the abelian case, the maximum is $(|A| + 1)/2$.) A set A whose doubling constant is maximum is called a *Sidon set*.

Another useful measure of additive structure between two sets $A, B \subseteq G$ is the *Ruzsa distance* d, defined by

$$d(A, B) = \log \frac{|A - B|}{\sqrt{|A|\,|B|}},$$

which satisfies all properties of a metric except that $d(A, A) \neq 0$ in general (indeed, $d(A, A) = \log \delta[A]$) [72]. The fact that d satisfies the triangle inequality follows from *Ruzsa's triangle inequality*: for all subsets $A, B, C \subseteq G$,

$$|A - C| \leq \frac{|A - B|\,|B - C|}{|B|}.$$

An immediate consequence of Ruzsa's triangle inequality is that

$$|A - A| \leq \frac{|A + A|^2}{|A|}$$

(taking $B = -A$ and $C = A$), which we can rewrite as

$$\delta[A] \leq \sigma[A]^2,$$

giving one relation between the doubling and difference constants [72].

1.2 The Additive Energy

Counting solutions to equations of the form $a+b=c+d$, where $a,b,c,d \in A$, gives a quantity useful in studying problems of this form. We define the *additive energy* of sets $A, B \subseteq G$ to be

$$E(A,B) := \#\{(a,a',b,b') \in A \times A \times B \times B : a+b=a'+b'\}.$$

Colloquially, we say a quadruple (a,a',b,b') satisfying this equation is a "collision," the idea being that the sum of a and b collides with the sum of a' and b'. It is easy to show that

$$|A|\,|B| \leq E(A) \leq |A|\,|B| \max(|A|\,,|B|).$$

The lower bound comes from assuming each pair $(a,b) \in A \times B$ yields a different sum. The upper bound comes from the fact that choosing a, b, and a' forces b' to equal $a+b-a'$ in order for the equation to be satisfied; likewise, choosing a, b, and b' forces our choice for a'.

The additive energy has analytic properties of great interest in additive combinatorics. Let G be a finite additive group. Define the Fourier transform of a function $f : G \to \mathbb{C}$ by

$$\widehat{f}(\chi) = \sum_{x \in G} f(x)\chi(x),$$

where χ is an additive character, and define the convolution of two functions $f, g : G \to \mathbb{C}$ by

$$f * g(x) = \sum_{t \in G} f(t)g(x-t).$$

Then $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$. We also have Parseval's identity:

$$\sum_{x \in G} |f(x)|^2 = \frac{1}{|G|} \sum_{\xi \in G} \left|\widehat{f}(\xi)\right|^2.$$

Let $f = \mathbf{1}_A$ and $g = \mathbf{1}_B$ be the indicator functions for the subsets $A, B \subseteq G$, and let $r(x) = \#\{(a,b) \in A \times B : a+b=s\}$ be the number of *representations* of x as a sum

of two elements of A. Then

$$r(x) = \sum_{t \in G} \mathbf{1}_A(t)\mathbf{1}_B(x-t) = \mathbf{1}_A * \mathbf{1}_B(x)$$

and

$$E(A,B) = \sum_{x \in G} r(x)^2 = \sum_{x \in G} \left(\mathbf{1}_A * \mathbf{1}_B(x)\right)^2 = ||\mathbf{1}_A * \mathbf{1}_B||_2^2 .$$

Taking $A = B$ and using Parseval's identity we have

$$E(A,A) = \frac{1}{|G|} \sum_{\xi \in G} \left|\widehat{\mathbf{1}_A}(\xi)\right|^4 = \frac{1}{|G|} \left|\left|\widehat{\mathbf{1}_A}\right|\right|_4^4 .$$

We also have the following identities:

$$|A| = \widehat{\mathbf{1}_A}(0) = \frac{1}{|G|} \sum_{\xi \in G} \left|\widehat{\mathbf{1}_A}(\xi)\right|^2 = \frac{1}{|G|} \left|\left|\widehat{\mathbf{1}_A}\right|\right|_2^2 .$$

Analyzing the Fourier coefficients at nonzero frequencies of the indicator function of A tells us a great deal about the additive structure (or lack thereof) in A. For example, if the nonzero Fourier coefficients of $\mathbf{1}_A$ are all sufficiently small, A is "pseudorandom" enough to contain a three-term arithmetic progression, and if there is a large nonzero Fourier coefficient, then A has dense intersection with a large arithmetic progression. This idea forms the basis of one proof of Roth's theorem on three-term arithmetic progressions [72].

1.3 Growth Theorems

Intuitively, we may expect a set with low additive energy (close to $|A|^2$) to have large growth in its sum set and a set with large additive energy (close to $|A|^3$) to have small growth. While the former statement is true, the latter statement is not.

By an application of the Cauchy-Schwarz inequality, we have

$$E(A) = \sum_{s \in A+A} r(s)^2 \geq \frac{1}{|A+A|} \left(\sum_{s \in A+A} r(s) \right)^2 = \frac{|A|^4}{|A+A|}.$$

Thus, $E(A) \geq \frac{|A|^4}{|A+A|}$. As a consequence, if $E(A)$ is asymptotically smaller than $|A|^3$, we obtain a nontrivial lower bound on the size of $A + A$.

We do not obtain any information from this bound if $E(A)$ is asymptotically equal to $|A|^3$. For example, suppose A is the disjoint union of the sets A_1 and A_2, where $|A_1| = |A_2| = n$, $|A_1 + A_1| = \Theta(n)$, and $|A_2 + A_2| = \Theta(n^2)$ (we shall present examples of sets A_1 and A_2 satisfying these properties in the following section). Since $A_2 + A_2 \subseteq A + A$, $|A + A| = \Theta(n^2)$, but it is also true that $E(A) \geq E(A_1) + E(A_2)$ (since the set of collisions within A will contain the set of collisions within A_1 and the set of collisions within A_2). Thus,

$$E(A) \geq \frac{n^4}{c_1 n} + \frac{n^4}{c_2 n^2} = \Theta(n^3).$$

So A is a set with asymptotically maximal growth and asymptotically maximal energy. On the other hand, the set A_1 has asymptotically minimal growth yet asymptotically maximal energy.

A "best-possible" result describing the structure of a set A with asymptotically maximal additive energy comes from a theorem of Antal Balog and Szemerédi, stating that there is a dense subset of A that has small growth [5]. The theorem was later given a new proof by Timothy Gowers [43].

Theorem 1 (Balog-Szemerédi-Gowers)**.** *Let A be a subset of an additive group. Given $c > 0$, there exist $c' = c'(c) > 0$ and $C = C(c) > 0$ such that, if $E(A, A) \geq c\,|A|^3$, then there is a subset $A' \subseteq A$ such that $|A'| \geq c'\,|A|$ and $|A' + A'| \leq C\,|A'|$.*

Gowers' proof considerably strengthened the bounds on the coefficients c' and C, giving them polynomial dependence (instead of tower-like dependence) on the parameter c. This means that the theorem still holds even if c is on the order of $|A|^{\varepsilon}$ for some fixed $\varepsilon > 0$ [72], a fact we shall use extensively later in the paper.

Another well-known and useful result proved via graph-theoretic methods is the Plünnecke-Ruzsa-Petridis inequality: [57, 58, 59, 56]

Theorem 2. *Let A be a finite subset of an additive group G such that $|A + A| \leq K\,|A|$. Then $|mA - nA| \leq K^{m+n}\,|A|$ for all $m, n \geq 1$.*

This theorem states that sets which are "almost closed" under the addition operation remain "almost closed" under multiple iterations of addition and subtraction. A similar result holds for multiplicative sets, as well:

Corollary 3. *Let A be a finite subset of the units of the ring R such that $|A.A| \leq K|A|$. Then $\left|A^{(m)}/A^{(n)}\right| \leq K^{m+n}|A|$ for all $m, n \geq 1$.*

1.4 The Erdős-Szemerédi Conjecture

A motivating example for the sum-product problem is the difference in behavior between different types of progressions in a ring. A *progression* is a set P in a semigroup $(G, +)$ of the form

$$P = \{a + kd : 0 \leq k \leq n\}$$

for some elements $a, d \in G$ and $n \in \mathbb{N}$. If the operation on G is written additively, then P is called an *arithmetic progression*, and if the operation is written multiplicatively, then P is called a *geometric progression*. In a ring, therefore, both types of progressions exist.

Consider the sets

$$P_1 = \{1, 2, 3, ..., n\}$$

and

$$P_2 = \{2, 4, 8, ..., 2^n\}$$

in the ring $\mathbb{Z}$. Both sets are progressions of size n, and P_1 is an arithmetic progression while P_2 is a geometric progression. Their sum sets are

$$P_1 + P_1 = \{2, 3, 4, 5, ..., 2n\},$$

an arithmetic progression of size $2n - 1$, while

$$P_2 + P_2 = \{2^i + 2^j : 1 \leq i, j \leq n\}$$

has size $\frac{n^2+n}{2}$. (To see this, rewrite the elements of P_2 in binary notation, and then observe that adding two distinct elements of P_2 results in a number with 1s in two different binary positions, while adding two of the same element of P_2 results in a number with a single 1.) Therefore, the sum set of the arithmetic progression only grows in size by a factor of $\Theta(1)$, while the sum set of the geometric progression grows by a factor of $\Theta(n)$.

Very different behavior can be seen in the product sets of these progressions. We have

$$P_2.P_2 = \{2^2, 2^3, 2^4, 2^5, ..., 2^{2n}\},$$

which has size $2n-1$, but $P_1.P_1$ has the size of the number of distinct elements in an $n \times n$ multiplication table, which was shown to be $\Theta(n^2/\log(n)^\gamma \log\log(n)^{3/2})$ for a constant γ by Kevin Ford [37]. Therefore, the product set of a geometric progression shows growth by a factor of $\Theta(1)$, and the product set of an arithmetic progression shows growth by nearly a factor of n.

These progressions display the extremes of additive and multiplicative structure in $\mathbb{Z}$. It is also rather easy to find a set that has large growth under both additive and multiplicative doubling: take P_1 to be the same as above and

$$P_2 = \{(3n)^k : 1 \leq k \leq n\},$$

and then define $A = P_1 \cup P_2$. Then $|A + A| = \Theta(n^2)$, since adding elements of P_1 and P_2 will yield unique sums, and $|A.A| = \Theta(n^2)$ as well, since multiplying elements of P_1 and P_2 will yield unique products. Therefore, both the sum set and product set of A have maximal growth.

Is it possible to find a set that has less-than-maximal growth in both its sum set and its product set? This is a much more difficult question, and one that remains open to this day. The question was first addressed in the literature by Paul Erdős and Endre Szemerédi, who proved the following theorem in the ring $\mathbb{Z}$:

Theorem 4 ([34]). *There is an absolute constant c such that, for every finite set A of integers,*

$$\max(|A+A|, |A.A|) \geq |A|^{1+c}.$$

In other words, no set of integers can have both small additive growth and small multiplicative growth. They further conjectured that

Conjecture 5 ([34]). *For every $\varepsilon > 0$, there exists an n_0 such that for every set of integers A with $|A| > n_0$,*

$$\max(|A+A|, |A.A|) \geq |A|^{2-\varepsilon}.$$

In other words, every sufficiently large set of integers must have nearly maximal growth in either its sum set or its product set; a sufficiently large set of integers cannot have both additive and multiplicative structure.

Erdős and Szemerédi did not attempt to give an explicit value for their constant c in Theorem 4, but explicit values were found in rapid succession by Melvyn Nathanson (who gave $c = 1/31$ in [54]), Kevin Ford (who gave $c = 1/15$ in the ring of real numbers $\mathbb{R}$ [36]), and György Elekes (who gave $c = 1/4$ in $\mathbb{R}$ [28]).

Elekes's proof is notable in that it obtains a much stronger bound than previously known in just two pages of work; it is also notable for being one of the first applications of incidence geometry to the sum-product problem. The proof invokes a theorem of Szemerédi and William T. Trotter:

Theorem 6 ([68, 67]). *There exists an absolute constant c such that the number of incidences between n points and t lines in $\mathbb{R}^2$ is at most $c(n^{2/3}t^{2/3} + n + t)$.*

Given a finite set of real numbers A, Elekes considered the set of $|A|^2$ lines in $\mathbb{R}^2$ of the form $y = a(x - b)$, each of which intersects the cartesian grid of points $(A+A) \times (A.A)$ exactly $|A|$ times. Employing a simple corollary of Theorem 6—if P

is a set of n points in $\mathbb{R}^2$ and $k \geq 2$, the number of lines intersecting at least k points of P is at most $C(\frac{n^2}{k^3} + \frac{n}{k})$—he showed that

$$|A|^2 \leq C\frac{|A+A|^2 |A.A|^2}{|A|^3},$$

yielding $|A+A||A.A| \geq |A|^{5/2}$ (up to a constant factor), whence the result.

This remarkably simple and effective technique was a breakthrough in the field, leading to new approaches on the sum-product problem using methods from incidence geometry. The most successful application of this approach to date was found by Józef Solymosi, who proved that $|A+A|^2 |A.A| = \Omega(|A|^4 / \log|A|)$; in other words, $\max(|A+A|, |A.A|) = \Omega(|A|^{4/3} / \log|A|)$, meaning that the constant c in Theorem 4 may be taken to be arbitrarily close to 1/3 [64].

1.5 Small Additive Growth and Small Multiplicative Growth

A problem related to the Erdős-Szemerédi problem is the study of the behavior of sets that have small additive or multiplicative growth. Elekes and Ruzsa [32] showed that if $|A+A| \leq |A|^{1+\varepsilon}$ for $\varepsilon > 0$, then $|A.A| = \Omega(|A|^{2-4\varepsilon} / \log|A|)$. Solymosi's aforementioned bound in [64] implies $|A.A| = \Omega(|A|^{2-2\varepsilon} / \log|A|)$ if $|A+A| \leq |A|^{1+\varepsilon}$. Mei-Chu Chang gives a more general result in the real numbers in return for a stronger hypothesis in [22]: if $|A+A| < K|A|$ with $K < K(\varepsilon, j, |A|)$ (which she remarks must be $o_{\varepsilon,j}(\log\log|A|)$), then $\left|A^{(j)}\right| = \Omega(|A|^{j-\varepsilon})$.

Progress on the corresponding result for when the product set is small has been slower. Solymosi's result above (as well as earlier results [31]) implies that if $|A.A| \leq |A|^{1+\varepsilon}$, then $|A+A| = \Omega(|A|^{3/2-\varepsilon})$. In [20], Mei-Chu Chang used methods based on Freiman's theorem to show that for $A \subseteq \mathbb{Z}$, if $|A.A| \leq c|A|$ for $c > 0$, then there exists a constant $c' = c'(c)$ such that $|A+A| \geq c'|A|^2$. Later in [22], Chang improved her result in the integers: if $|A.A| \leq |A|^{1+\varepsilon}$, then $|jA| \geq |A|^{j-\delta_j(\varepsilon)}$, where $\delta_j(\varepsilon) \to 0$ as $\varepsilon \to 0$. In the same paper, she utilized a version of Freiman's theorem—

Lemma 7 ([39]). *If G is a torsion-free abelian group, $A \subseteq G$, and $|A.A| < K|A|$, then*

$$A \subseteq \{g_1^{j_1} \cdots g_d^{j_d} : j_i = 1, ..., \ell_i, \text{ and } g_i \in G\},$$

where $d \leq K$ and $\prod \ell_i < c(K)|A|$.

—and the subspace theorem—

Theorem 8 ([35, 2, 61]). *Let k be an algebraically closed field, let Γ be a multiplicative subgroup of k^* of rank r, and let $a_1, a_2, ..., a_n \in k^*$. Then there are at most $(8n)^{4n^4(n+nr+1)}$ solutions $(z_1, ..., z_n) \in \Gamma^n$ to the equation*

$$a_1 z_1 + a_2 z_2 + \cdots + a_n z_n = 1$$

with no vanishing subsum on the left-hand side.

—to show that if A is a subset of the *real* numbers with $|A.A| < K|A|$, then for every $\varepsilon > 0$, $|jA| > |A|^{j-\varepsilon}$ provided $K = o_{j,\varepsilon}(\log |A|)$. However, as noted, because these results rely on Freiman's theorem, they cannot obtain nontrivial results in $\mathbb{R}$ or $\mathbb{C}$ when $|A.A| < K|A|$ for $K = |A|^{\varepsilon}$, $\varepsilon > 0$.

Solymosi proved an Elekes-like bound ($c = 1/4$) for sets A in the complex numbers, quaternions, and other hypercomplex numbers [62]. In 2013, Sergei Konyagin and Misha Rudnev improved this bound for $\mathbb{C}$ to match Solymosi's bound for $\mathbb{R}$ ($c = 1/3$) [53].

1.6 The Sum-Product Conjecture in Fields of Prime Order

The sum-product problem can also be studied in rings and fields that do not properly contain the integers. One natural setting in which to study the problem is in finite fields. It is known that it is not possible to obtain the result analogous to the Erdős-Szemerédi conjecture (that is, $\max(|A + A|, |A.A|) \geq c(\varepsilon) \min(|A|^{2-\varepsilon}, q^{1-\varepsilon})$ for all $A \subseteq \mathbb{Z}/q\mathbb{Z}$) [40].

Jean Bourgain, Netz Katz, and Terence Tao provided one of the first major sum-product results in this setting:

Theorem 9 ([16]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q, and let $A \subset F$ such that $q^\delta < |A| < q^{1-\delta}$ for some $\delta > 0$. Then there exist $c = c(\delta) > 0, \varepsilon = \varepsilon(\delta) > 0$ such that $\max(|A+A|, |A.A|) \geq c\,|A|^{1+\varepsilon}$.*

The authors then applied Theorem 9 to deduce a Szemerédi-Trotter theorem for finite fields:

Theorem 10 ([16]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q, and let $\mathbb{P}_F^3$ be the projective plane over F. If $P \subset \mathbb{P}_F^3$ is a set of points and L is a set of lines in $\mathbb{P}_F^3$ such that $|P|, |L| \leq N = |F|^\alpha$ for $0 < \alpha < 2$, then there exists $\varepsilon = \varepsilon(\alpha) > 0$ such that*

$$\#\{(p, \ell) \in P \times L : p \in \ell\} \leq CN^{3/2-\varepsilon}.$$

This Szemerédi-Trotter-type result further permits one to obtain a bound on the Erdős distance problem in prime fields in which -1 is not a quadratic residue:

Theorem 11 ([16]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime $q \equiv 3 \pmod 4$, and let $P \subseteq F^2$ be a set of size $N = |F|^\alpha$ for some $0 < \alpha < 2$. Then there exists $\varepsilon = \varepsilon(\alpha) > 0$ such that*

$$\#\{d(p, p') : p, p' \in P\} \geq CN^{1/2+\varepsilon},$$

where $d((x_1, y_1), (x_2, y_2)) = (x_1 - x_2)^2 + (y_1 - y_2)^2$.

Shortly after [16], Bourgain along with Alexei Glibichuk and Sergei Konyagin published a result removing the need for a lower bound on $|A|$ in Theorem 9:

Theorem 12 ([16]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q, and let $A \subset F$ such that $|A| < q^{1/2}$. Then there exist $c > 0, \varepsilon > 0$ such that $\max(|A+A|, |A.A|) \geq c\,|A|^{1+\varepsilon}$.*

In the same paper, Bourgain, Glibichuk, and Konyagin gave an estimate on exponential sums stating that multiplicative subgroups of $\mathbb{F}_p^*$ have little additive structure [44]:

Theorem 13 ([15]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q, let $\delta > 0$, and let $H \leq F^*$ be a multiplicative subgroup of F^* of size at least q^δ. Then there exists $\varepsilon = \varepsilon(\delta) > 0$ such that for all $\xi \neq 0$,*

$$\frac{1}{|H|}\left|\sum_{x\in H} e^{2\pi i x\xi/q}\right| \leq Cq^{-\varepsilon}.$$

In 2007 Derrick Hart, Alex Iosevich, and Solymosi gave the following explicit estimates using Kloosterman sums.

Theorem 14 ([46]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q. Then*

$$\max(|A+A|, |A.A|) \geq \begin{cases} C|A|^{3/2} q^{-1/4}, & q^{1/2} < |A| < q^{7/10} \\ C|A|^{2/3} q^{1/3}, & q^{7/10} < |A| \leq q. \end{cases}$$

In 2007 and 2008, Moubariz Garaev published two explicit estimates for prime fields when $|A|$ is smaller than $q^{7/13}$ and when $|A|$ is larger than $q^{2/3}$:

Theorem 15 ([40, 41]). *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q. Then*

$$\max(|A+A|, |A.A|) \geq \begin{cases} C\frac{|A|^{5/3}}{q^{1/3}\log|A|}, & |A| \leq q^{7/13}(\log q)^{-4/13} \\ Cq^{1/2}|A|^{1/2}, & |A| > q^{2/3}. \end{cases}$$

The latter situation is optimal for sets of size greater than $q^{2/3}$, since it is possible to construct a subset $A \subseteq \mathbb{Z}/q\mathbb{Z}$ of any size such that $\max(|A+A|, |A.A|) \leq cq^{1/2}|A|^{1/2}$ [41]. Solymosi gives a different proof of Garaev's bound for $|A| > q^{2/3}$ using expander graphs [65].

Katz and Chun-Yen Shen modified Garaev's argument in [40] to improve his result for $|A| < q^{1/2}$ to

$$\max(|A+A|, |A.A|) \geq |A|^{14/13}.$$

1.7 Other Variants of the Sum-Product Problem

A variant of the sum-product problem is to replace the set $A + A$ or $A.A$ (or both) with different "functions" in the set A: more precisely, sets of the form

$$f(A,A) = \{f(a,b) : (a,b) \in A \times A\}$$

(or analogously for functions with other than two variables). Bourgain published a result of this flavor in [9], proving a sum-product-type inequality for $A + A$ versus $1/A + 1/A$. Elekes also showed results of this type: if f is a strictly convex or concave function and $A \subseteq \mathbb{R}$, then $|A \pm A|\,|f(A) \pm f(A)| = \Omega(|A|^{5/2})$(giving the inequality $\max(|A + A|, |1/A + 1/A|) = \Omega(|A|^{5/4})$) and $|A + 1/A| = \Omega(|A|^{5/4})$ [30], resolving a question of Erdős and Szemerédi in [34].

Van Vu studied this problem for polynomial functions in finite fields and proved the following result:

Theorem 16 ([73, 19])**.** *Let $F = \mathbb{Z}/q\mathbb{Z}$ for some prime q. There is an absolute constant $c > 0$ such that, if f is a polynomial of degree d in $F[x, y]$ and not of the form $g(\ell(x, y))$ for a polynomial $g \in F[t]$ and a linear form $\ell \in F[x, y]$, then for all $A \subseteq F$ with $|A| > \sqrt{q}$,*

$$\max(|A + A|, |f(A, A)|) \geq c\,|A| \begin{cases} (|A|/\sqrt{q})^{1/2} d^{-2}, & \sqrt{q} < |A| \leq d^{4/5} q^{7/10} \\ (q/|A|)^{1/3} d^{-1/3}, & |A| \geq d^{4/5} p^{7/10}. \end{cases}$$

The theorem immediately implies the previously mentioned sum-product result of Hart, Iosevich, and Solymosi in $\mathbb{Z}/q\mathbb{Z}$ by taking $f(x, y) = xy$ [46, 73].

In [19] Boris Bukh and Jacob Tsimerman gave sum-product-type inequalities for polynomial functions for small subsets of $F = \mathbb{Z}/q\mathbb{Z}$, including an extension of Vu's result to sets with $|A| < \sqrt{p}$ for quadratic polynomials. They further developed sum-product-type inequalities for rational functions of large subsets of finite fields: for example, if $f \in F(x)$ and $g \in F(x, y)$ are nonconstant rational functions of degree $d < q^{1/50}$ and g is not of the degenerate form $G(af(x) + bf(y) + c)$, $G(x)$, or $G(y)$, then for all $A \subseteq F$ with $|A| \geq \sqrt{q}$, we have

$$\max(|f(A) + f(A)|, |g(A, A)|) \geq c\,|A| \begin{cases} (|A|/\sqrt{q})^{1/2} d^{-2}, & \sqrt{q} \leq |A| \leq d^{8/5} q^{7/10} \\ (q/|A|)^{1/3} d^{-2}, & |A| \geq d^{8/5} q^{7/10}. \end{cases}$$

Recently Terence Tao established that for a bivariate polynomial $f \in F[x, y]$ of bounded degree, with F a finite field of large characteristic, either $|f(A, B)| = \Theta(|F|)$ whenever A and B are subsets of F such that $|A|\,|B| = \Omega(|F|^{15/8})$ or else f has the form $f(x, y) = g(p(x) + q(y))$ or $g(p(x)q(y))$ for polynomials g, p, q [71]. His main new tool in attacking the problem is an algebraic version of the Szémeredi regularity lemma, which he uses to describe the structure of graphs generated by subsets of finite fields of large characteristic [71].

The sum-product problem can also be studied in matrix spaces. Mei-Chu Chang provided some early results in this setting [23, 72]:

Theorem 17. *There is a function $\Phi(n)$ tending to infinity with n such that, if d is a fixed integer and A is a finite set of $d \times d$ matrices with real entries such that $\det(M - M') \neq 0$ for all pairs of distinct matrices $M, M' \in A$, then*

$$|A + A| + |A.A| \geq \Phi(|A|)\,|A|\,.$$

Moreover, for every positive integer d there exists $\varepsilon = \varepsilon(d) > 0$ such that

$$|A + A| + |A.A| \geq |A|^{1+\varepsilon}\,.$$

Focusing on the multiplicative structure of matrix rings, Chang shows that if $A \subset \mathrm{SL}_3(\mathbb{Z})$ does not have large intersection with any cosets of a nilpotent subgroup, then $|A.A.A| \geq c\,|A|^{1+\varepsilon}$. Chang also proves a similar result for $\mathrm{SL}_2(\mathbb{C})$ [24, 70].

Chang also attains a sum-product result ($|A + A| + |A.A| \geq |A|^{1+c}$ for an absolute constant c) in semisimple commutative Banach algebras using Freiman's lemma and the Balog-Szemerédi-Gowers theorem [21]. However, in this setting, it is known that $c \leq 1 - \frac{\log 2}{\log 3}$ [70].

1.8 Connections to the Theory of Expanders

Let $G = G(V, E)$ be a (simple, undirected) graph, and let $N(W)$ denote the set of vertices in $V \setminus W$ which are adjacent to some vertex in W. We say G is an (n, d, c)-*expander* if it has n vertices, the maximum degree of a vertex is d, and for every set of vertices $W \subseteq V$, we have $|N(W)| \geq c\,|W|$. The study of families of expanders (that is, sequences $(G_1, G_2, ...)$ such that each G_i is an (n_i, d, c)-expander and $n_i \to \infty$ with i) has many applications to problems in theoretical computer science, such as circuit construction, error correcting codes, and complexity theory [1, 49].

Results in sum-product inequalities are closely connected to the theory of expanders. For example, the previously stated results by Solymosi and Vu on the sum-product and sum-polynomial problems in finite fields are proved through the study of the eigenvalues of expander graphs [65, 73]. Conversely, results on sum-product inequalities have implications in the theory of expanders.

Harald Helfgott obtained results of this flavor working towards the following conjecture of László Babai and Ákos Seress:

Conjecture 18 ([4])**.** *For a group G and set of generators A, let $\Gamma(G, A)$ be the Cayley graph (whose vertex set is G and whose edge set is $\{(ag, g) : g \in G, a \in A\}$). Let the diameter of a graph Γ be the maximum over all pairs (u, v) of vertices in Γ of the length of the shortest path between u and v.*

There is an absolute constant $c > 0$ such that for every nonabelian finite simple group G and set of generators A of G,

$$\operatorname{diam}(\Gamma(G, A)) = O((\log |G|)^c).$$

Helfgott proves the conjecture for $G = \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$, q prime, as a direct consequence of a growth theorem for multiplicative subsets of $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$:

Theorem 19 ([47])**.** *Let q be a prime, and let A be a subset of $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ not*

contained in any proper subgroup. For all $\delta > 0$, there exist $c = c(\delta) > 0$ and $\varepsilon = \varepsilon(\delta) > 0$ such that if $|A| < p^{3-\delta}$, then $|A.A.A| > c\,|A|^{1+\varepsilon}$.

If ψ is a symmetric probability distribution on G (that is, $\psi(g) = \psi(g^{-1})$ for all $g \in G$) whose support contains A, define the transition matrix $T_\psi(G, A) = (\psi(y^{-1}x))_{x,y\in G}$. Then the largest eigenvalue of $T_\psi(G, A)$ is 1; define the *spectral gap* of $T_\psi(G, A)$ to be the difference between 1 and the second-largest eigenvalue. An alternate definition of a family of expander graphs is to consider a family $\{G_j, A_j\}_{j\in J}$ of finite groups G_j and sets of generators A_j of G_j such that $d = \left|A_j \cup A_j^{-1}\right|$ is constant. Letting $\psi_j(g) = 1/d$ for $g \in A_j \cup A_j^{-1}$ and $\psi_j(g) = 0$ for $g \notin A_j \cup A_j^{-1}$, we may define $\{\Gamma(G_j, A_j)\}_{j\in J}$ to be a family of expander graphs if the spectral gap of $T_{\psi_j}(G_j, A_j)$ is bounded below by a positive constant [47].

Helfgott's result did not show that $\{\Gamma(G, A)\}$ (where G varies over $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ for all the primes q and A varies over all sets of generators such that $|A \cup A^{-1}|$ is fixed) is a family of expanders, but it did yield nontrivial consequences on the spectral gap of $T_\psi(G, A)$ and the mixing time of the Cayley graph $\Gamma(G, A)$:

Corollary 20 ([47])**.** *Let q be a prime and A be a set of generators for $G = \mathrm{SL}(\mathbb{Z}/q\mathbb{Z})$. Let ψ be a symmetric probability distribution on G whose support contains A, define the transition matrix $T_\psi(G, A) = (\psi(y^{-1}x)_{x,y\in G})$, and let $\eta = \min\{\psi(g) : g \in A \cup A^{-1}\}$.*

Then the second largest eigenvalue of $T_\psi(G, A)$ is at most $1 - C\eta^{-1}(\log p)^{-2c}$ for absolute constants $C, c > 0$.

Furthermore, the mixing time of $\Gamma(G, A)$ is at most $C\,|A|\,(\log p)^{2c+1})$ for absolute constants $C, c > 0$.

Later, Bourgain and Gamburd showed that if A is a set of generators of $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ such that the girth of $\Gamma(G, A)$ is $\Omega(\log |G|)$, then the adjacency matrix of $\Gamma(G, A)$ has spectral gap bounded below by a constant [48, 13].

A result of Gowers and Babai, Nikolov, and Pyber states that if $A \subseteq G = \mathrm{SL}_n(\mathbb{Z}/q\mathbb{Z})$ and $|A| > 2\,|G|^{1-1/(3n+3)}$, then $A.A.A = \mathrm{SL}_n(K)$ [48, 42, 55, 3].

Helfgott proved a later result for $\mathrm{SL}_3(\mathbb{Z}/q\mathbb{Z})$ using the sum-product theorem of Bourgain, Katz, Tao, and Konyagin:

Theorem 21 ([48])**.** *Let q be a prime. For all $\varepsilon > 0$, if A be a set of generators of $G = \mathrm{SL}_3(\mathbb{Z}/q\mathbb{Z})$ such that $|A| < |G|^{1-\varepsilon}$, there exist $\delta = \delta(\varepsilon) > 0$ and $c = c(\varepsilon) > 0$ such that*

$$|A.A.A| \geq c\,|A|^{1+\delta}.$$

This theorem, along with the result of Gowers and Babai, Nikolov, and Pyber, implies Babai's conjecture for $G = \mathrm{SL}_3(\mathbb{Z}/q\mathbb{Z})$ [48]. Later, Breuillard, Green, and Tao extended Helfgott's result to show that approximate subgroups of $\mathrm{SL}_n(\mathbb{F}_q)$ that generate the group are either very small or else make up nearly all of the group [18]. In contrast to Helfgott's work, Breuillard, Green, and Tao do not use sum-product theorems to obtain their result; instead, they derive the sum-product theorem from their result on approximate subgroups.

1.9 Summary of New Contributions

In this section we detail the contributions of the present work and their significance in the context of the present state of research.

1.9.1 Sets of Rich Lines in General Position

In this chapter we prove a δ-ε formulation of a conjecture of Solymosi initially published by Elekes [30]. This conjecture implies a statistical version of a Freiman-type theorem on linear functions with small image sets. The following exposition is primarily based on selected sections of Elekes' survey paper [30], and the reader is directed to the references and proofs therein.

Let G be an additive abelian group. A *generalized arithmetic progression*, or GAP, $P \subseteq G$, is a set of the form

$$P = \{a_0 + r_1 a_1 + \cdots + r_d a_d : 0 \leq r_1 \leq n_1, ..., 0 \leq r_d \leq n_d\},$$

where $a_0, ..., a_d \in G$, $n_1, ..., n_d \in \mathbb{Z}^+$, and $|P| = n_1 \cdots n_d$.[1] For multiplicative abelian groups G, an analogous definition may be made: a *generalized geometric progression*, or GGP, $P \subseteq G$, is a set of the form

$$P = \{a_0 \cdot r_1 a_1 \cdots r_d a_d : 0 \leq r_1 \leq n_1, ..., 0 \leq r_d \leq n_d\},$$

where $a_0, ..., a_d \in G$, $n_1, ..., n_d \in \mathbb{Z}^+$, and $|P| = n_1 \cdots n_d$. In both these cases, the *dimension* of P, $\dim(P)$, is defined to be d.

Freiman showed that a set $A \subset \mathbb{R}$ with a small sum set must be contained in a proportionally sized GAP and bounded dimension:

Theorem 22 ([38, 39])**.** *For all $c > 0$ there exist $C = C(c) > 0$ and $d = d(c) > 0$ such that the following property holds.*

Let $A \subset \mathbb{R}$. If $|A + A| \leq c\,|A|$, then A is contained in a GAP P such that $|P| \leq C\,|A|$ and $\dim(P) \leq d$.

This theorem also holds for multiplicative subsets of the nonzero reals, replacing $|A + A|$ with $|A.A|$ and GAP with GGP.

If $A \subset \mathbb{R}$, let $E \subseteq A \times A$. In this way, E can be considered the edge set of a bipartite $A \sqcup A$. We define the *restricted sum set*

$$A +_E A := \{a + b : (a, b) \in E\}.$$

Theorem 23 ([5])**.** *For all $a, c > 0$ there exist $\alpha = \alpha(a, c) > 0$, $C = C(a, c) > 0$, and $d = d(a, c) > 0$ such that the following property holds.*

If $|E| \geq a\,|A|^2$ and $|A +_E A| \leq c\,|A|$, then there is a subset $A' \subseteq A$ with $|A'| \geq \alpha\,|A|$ such that $|A' + A'| \leq C\,|A'|$.

[1] Some authors allow $|P| < n_1 \cdots n_d$ and call GAPs such that $|P| = n_1 \cdots n_d$ *proper*.

Combining Theorem 22 with Theorem 23, we see that a set A with small restricted sum set must intersect a GAP P with with $\dim(P) \leq d$ in at least $|P|/C$ places. This may be considered a "statistical" version of Freiman's theorem [30]. A "uniform statistical" hypothesis (that is, requiring a minimum degree proportional to $|A|$ in the bipartite graph G) guarantees that A can be covered by a constant number of GAPs [30].

Theorem 24 ([30, 33])**.** *For all $\delta, c, \varepsilon > 0$ there exist $C = C(\delta, c, \varepsilon) > 0$, $d = d(\delta, c, \varepsilon) > 0$, and $\gamma = \gamma(\delta, c)$ such that the following property holds.*

Let $A \subset \mathbb{R}$ be finite, G a bipartite graph on $A \sqcup A$ with minimum degree $\delta |A|$. If $|A +_E A| \leq c |A|$, then A can be partitioned into k disjoint sets $A_1, ..., A_k$ such that:

1. *Each A_i is contained in a GAP P_i with $|P_i| \leq C |A|$ and $\dim(P_i) \leq d$ (the GAPs P_i need not be disjoint);*
2. *For each i, at least $\gamma |A|^2$ edges of E are between elements of A_i (hence $k \leq 1/\gamma$); and*
3. *There are at most $\varepsilon |A|^2$ "leftover" edges. That is,*

$$\sum_{i<j} \sum_{a \in A_i} \sum_{b \in A_j} \mathbf{1}_E((a,b)) \leq \varepsilon |A|^2 .$$

It turns out that Freiman's theorem extends to more general objects than sets of real numbers: sets of linear functions [30]. Sets of lines with small composition sets satisfy a particular "two extremities" structure.

For the remainder of the section, let $\mathcal{L}$ be the set of all functions $f : \mathbb{R} \to \mathbb{R}$ of the form $f(x) = mx + b$ for real numbers $m \neq 0$ and b. If $L \subseteq \mathcal{L}$, define $L^{-1} = \{f^{-1} : f \in L\}$.

If $L \subseteq \mathcal{L}$, $P \subseteq L$ is a *parallel family* if the graphs of the lines in P are parallel, and $S \subseteq L$ is a *star family* if the graphs of the lines in S intersect at some common point. We say L is in *general position* if L has no parallel families of size two and

no star families of size three. If L has no parallel and no star families of size greater than some constant $C > 2$, then we say L is in *near-general position*.

Theorem 25 (Elekes [27, 30])**.** *For all $c, C > 0$, there exists $c' = c'(c, C) > 0$ such that the following property holds.*

Let $L_1, L_2 \subset \mathcal{L}$ be sets of n lines each and $E \subset L_1 \times L_2$ have size at least cn^2. Define

$$L_1 \circ_E L_2 := \{f \circ g : (f, g) \in E\}.$$

If $|L_1 \circ_E L_2| \leq Cn$, then there exist subsets $L_1' \subseteq L_1$ and $L_2' \subseteq L_2$ such that $|(L_1' \times L_2') \cap E| \geq c'n^2$ and both L_1' and L_2' are either parallel families (possibly of different slopes) or star families (possibly of different common intersections).

Using this theorem along with Theorem 22, Elekes proved a result analogous to Theorem 22—in fact, a true generalization of Freiman's theorem.

Theorem 26 (Elekes [29, 30])**.** *For every $c > 0$ there exist $C = C(c) > 0$ and $C' = C'(c) > 0$ such that the following property holds.*

If $L_1, L_2 \subset \mathcal{L}$ are sets of n lines each and $\left|L_1^{-1} \circ L_2\right| \leq cn$, then $L_1 \cup L_2$ is contained in a union of either C parallel families or C star families, and each of those families has size at most $C'n$.

A uniform statistical version of this Freiman-type theorem for linear functions can be deduced from Theorem 24:

Theorem 27 (Elekes [**?**])**.** *For all $\delta, c > 0$ there exists $C = C(\delta, c) > 0$ such that the following property holds.*

Let L_1, L_2 be as in Theorem 26, $G = (L_1 \sqcup L_2, E)$ a bipartite graph with minimum degree at least δn. If $\left|L_1^{-1} \circ_E L_2\right| \leq cn$, then $L_1 \cup L_2$ is the union of C parallel and star families.

The previous theorems have focused on sets of lines whose composition set is small. Another direction we can explore is to study sets of lines L and points $A \subset \mathbb{R}$ whose *image set* $L(A) = \{f(a) : f \in L, a \in A\}$ is small. In particular, Theorem 25 is equivalent to the following:

Theorem 28 (Elekes [30])**.** *For all $c, c' > 0$, there exists $C = C(c, c') > 0$ such that the following property holds.*

Let $L \subset \mathcal{L}$ and $A \subset \mathbb{R}$ each have size n, let $E \subset L \times A$ have size at least cn^2, and define

$$L_E(A) := \{f(a) : (f, a) \in E\}.$$

If $|L_E(A)| \leq c'n$, then there exists a parallel or star family $L' \subseteq L$ such that

$$|E \cap (L' \times A)| \geq Cn^2.$$

A result about small image sets analogous to Theorem 26 holds:

Theorem 29 (Elekes [30])**.** *For all $c > 0$ there exist $C = C(c) > 0$ and $C' = C'(c) > 0$ such that the following property holds.*

Let $L \subset \mathcal{L}$ and $A \subset \mathbb{R}$ each have size n. If $|L(A)| \leq cn$, then L is contained in the union of at most C parallel families or of at most C star families, and each of these families has size at most $C'n$.

Elekes asked whether a uniform statistical version of this theorem similar to Theorem 27 holds. We formulate the following conjecture in which a minimum degree is required on only one side of the bipartite graph.

Conjecture 30. *For all $\delta, c > 0$ there exists $C = C(\delta, c) > 0$ such that the following property holds.*

Let L, A be as in Theorem 28, $G = (L \sqcup A, E)$ a bipartite graph with degree at least δn for each $f \in L$. If $|L_E(A)| \leq cn$, then L is the union of C parallel and star families.

In terms of cartesian products, this conjecture is equivalent to the following:

Conjecture 31 (Elekes [30])**.** *If L is a set of cn lines, each cn-rich in an $n \times n$ cartesian product, then L is the union of $C = C(c)$ parallel and star families.*

Proposition 32. *Conjecture 30 and Conjecture 31 are equivalent.*

The following proof is not given explicitly in [30] but can be inferred from similar arguments presented in the paper.

Proof. Suppose Conjecture 30 holds. Fix $0 < c < 1$, and let L be a set of cn lines, each cn-rich in a $n \times n$ Cartesian product $A \times B$. Construct the bipartite graph $G = (L \sqcup A, E)$, where an edge connects $f \in L$ and $a \in A$ whenever $\phi(a) \in B$. Then the degree of each $f \in L$ is at least cn, so L is the union of a constant number of parallel and star families.

For the reverse implication, suppose L is a set of n lines, A is a set of n points, and the edge set $E \subseteq L \times A$ satisfies $|L_E(A)| \leq Cn$ for $C > 0$. Observe that each line from L occurs in at least δn pairs of E, so we have $n \geq \delta n$ lines which are each δn-rich in the cartesian product $(A \cup L_E(A)) \times (A \cup L_E(A))$, which has size at most $(C+1)^2 n^2$. So the lines are the union of a constant number of parallel and star families. □

Conjecture 31, and therefore Conjecture 30, would be implied by following:

Conjecture 33 (Solymosi [30])**.** *Among the lines which are cn-rich in an $n \times n$ cartesian product, at most $C = C(c)$ can be in general position.*

Proposition 34. *Conjecture 33 implies Conjecture 31*

Proof. Given a set L of cn lines which are cn-rich in an $n \times n$ grid, let L' be a maximum collection of these lines in general position. By Conjecture 33, $|L'| \leq C$. Define $L(\lambda)$ to be the set of lines in L with slope λ and define $L(p)$ to be the set of lines in L passing through a given point $p \in \mathbb{R}^2$. Then the union of $L(\lambda)$ over all λ

which are slopes of lines in L' and $L(p)$ over all points p which are intersections of pairs of lines in L' must be L. So L is the union of at most $C + C^2$ parallel and star families. □

The main result of Chapter II, Theorem 39, yields a version of Conjecture 31.

Corollary 35. *For every $\varepsilon > 0$, there exists $0 < \delta_0 < \varepsilon$ such that for all $0 < \delta < \delta_0$ and sufficiently large n, the following property holds.*

If each of $n^{1-\delta}$ lines is $n^{1-\delta}$-rich in an $n \times n$ cartesian product, then the set of lines is the union of n^{ε} parallel and star families.

Therefore, we also have the following uniform statistical Freiman-type theorem similar to Conjecture 30:

Corollary 36. *For every $\varepsilon > 0$, there exists $0 < \delta_0 < \varepsilon$ such that for all $0 < \delta < \delta_0$ and sufficiently large n, the following property holds.*

Let L, A be as in Theorem 28, $G = (L \sqcup A, E)$ a bipartite graph with degree at least $n^{1-\delta}$ for each $f \in L$. If $|L_E(A)| \leq n^{1+\delta}$, then L is the union of n^{ε} parallel and star families.

Proof. Let $B = A \cup L_E(A)$. Each line in L is incident with at least $n^{1-\delta}$ edges of E, so each line in L is $n^{1-\delta}$-rich in the $n^{1+\delta} \times n^{1+\delta}$ cartesian product $B \times B$. In other words, each line of L is $|B|^{1-\delta'}$-rich in $B \times B$, where $\delta' = 1 - \frac{1-\delta}{1+\delta} \to 0$ as $\delta \to 0$. By choosing δ small enough, we can ensure through Corollary 35 that L is the union of n^{ε} parallel and star families. □

1.9.2 Few Products, Many Differences

In this chapter, we give a short proof of a result similar in flavor to Solymosi's result [64] that a set A with a small sum set has many products. However, our result gives a converse bound: if A has a small product set, then $A - A$ has many elements. The proof is conditional on the following conjecture:

Conjecture 37. *Let $V \subset \mathbb{R}^2$ be a set of vectors such that at most $\frac{|V|}{2}$ are contained in any common line. Then*

$$\sum_{s>0} \#\{(\mathbf{v}, \mathbf{w}) \in V \times V : |\mathbf{v} \times \mathbf{w}| = s\}^2 = O(|V|^3 \log |V|)$$

(In Conjecture 37, $\mathbf{v} \times \mathbf{w}$ is the cross product of $\mathbf{v}$ and $\mathbf{w}$ considered as vectors in $\mathbb{R}^3$.)

One reason that Conjecture 37 is of interest is that it quickly implies a result about areas of triangles formed from a set of N points in the plane, which we suspect is also true:

Theorem 38 ([50])**.** *If Conjecture 37 holds, then there exists an absolute constant $c > 0$ such that $N > 1$ points in $\mathbb{R}^2$, not all on the same line, determine at least $cN/\log N$ areas of triangles with one vertex at the origin.*

The conjecture would also imply a sum-product-type inequality of the form

$$|A.A \pm A.A| \geq C(\varepsilon) |A|^{2-\varepsilon}$$

for all $\varepsilon > 0$ [50], a statement which we also suspect to be true.

1.9.3 Contributions of the Author

Sets of Rich Lines in General Position was primarily written by the author under the supervision of Prof. Ernie Croot with contributions from Albert Bush and Gagik Amirkhanyan. The latter two sections of the chapter were originally written by Bush and later revised and updated by the author.

The text of *Few Products, Many Differences* was originally written by Bush and later revised and updated by the author. The initial argument behind the proof of the chapter's main result was conceived jointly by the author and Bush; Amirkhanyan later strengthened the original argument to its present form. This chapter was also written under the supervision of Prof. Croot.

CHAPTER II

SETS OF RICH LINES IN GENERAL POSITION

2.1 Introduction

Our goal in the present chapter is to prove a δ-ε formulation of a conjecture of Solymosi found in [30]:

Theorem 39. *For every $\varepsilon > 0$, there exists $0 < \delta_0 < \varepsilon$ such that for all $0 < \delta < \delta_0$ and for sufficiently large $n = n(\varepsilon, \delta)$, the following property holds.*

If $A \subseteq \mathbb{R}$ has size n, then every set of at least n^{ε} lines in $\mathbb{R}^2$, each of which intersects $A \times A$ in at least $n^{1-\delta}$ points, contains either two parallel lines or three lines with a common intersection point.

The Szemerédi-Trotter theorem gives a bound of $O(n^{1+3\delta})$ $n^{1-\delta}$-rich lines for an arbitrary set of n^2 points. In addition to the connections Theorem 39 has to Freiman's theorem, a consequence of Solymosi's conjecture is that requiring a grid structure in the set of points and general position in the set of lines gives a significant improvement to the Szemerédi-Trotter bound.

We begin by listing the major tools and basic results needed to prove Theorem 39. Next we introduce a key result, Theorem 44, describing the behavior of sets of rich lines under self-composition. Theorem 44 will allow us to extract a set of lines in "nearly" general position from the composition of a set of rich lines with itself. We then prove a weakened version of Solymosi's conjecture:

Theorem 40. *For every $\varepsilon > 0$, there exists $0 < \delta_0 < \varepsilon$ such that for all $0 < \delta < \delta_0$ and for sufficiently large $n = n(\varepsilon, \delta)$, the following property holds.*

If $A \subseteq \mathbb{R}$, $|A| = n$, then every set of at least $n^{1-\varepsilon}$ lines in $\mathbb{R}^2$, each of which

intersects $A \times A$ in at least $n^{1-\delta}$ points, contains either two parallel lines or $C = C(\varepsilon) \geq 2$ lines with a common intersection point.

Finally, we use the subset-extraction theorem along with Theorem 40 to conclude the proof of Theorem 39.

2.2 Preliminaries

Let $A \subset \mathbb{R}$ be a finite subset with $|A| = n$. We call a line ℓ in $\mathbb{R}^2$ *k-rich* if it intersects at least k points in $A \times A$. A line can be at most n-rich; we will concern ourselves mainly with lines that are $n^{1-\delta}$-rich for some small positive δ.

If $\ell : y = \lambda x + b$ is a line in $\mathbb{R}^2$, then $\ell^{-1} : y = \frac{1}{\lambda}x - \frac{b}{\lambda}$ is the line such that $\ell \circ \ell^{-1} = \ell^{-1} \circ \ell$ is the identity function on $\mathbb{R}^2$ (i.e., the line $y = x$).

In the next two lemmas, we establish that many pairs of lines in a set of rich lines can be combined to obtain lines of slightly less richness.

Lemma 41. *Given sets $A_1, ..., A_k \subseteq \{1, ..., n\}$, each of size at least $n^{1-\delta}$, we must have at least $k^2 n^{-2\delta}/2$ ordered pairs of sets (A_i, A_j) with $|A_i \cap A_j| \geq n^{1-2\delta}/2$.*

Proof. Let $B = \{(i,j) : |A_i \cap A_j| \geq \frac{n^{1-2\delta}}{2}\}$. For a contradiction, suppose $|B| < \frac{1}{2}k^2 n^{-2\delta}$. Then

$$\sum_{i,j} |A_i \cap A_j| = \sum_{(i,j)\in B} |A_i \cap A_j| + \sum_{(i,j)\in B^c} |A_i \cap A_j| < n \cdot \frac{1}{2}k^2 n^{-2\delta} + \left(k^2 - \frac{1}{2}k^2 n^{-2\delta}\right)\frac{n^{1-2\delta}}{2} < k^2 n^{1-2\delta}.$$

However, letting $d(x) := \#\{1 \leq i \leq k : x \in A_i\}$, we have by Cauchy-Schwarz

$$\sum_{i,j} |A_i \cap A_j| = \sum_{x=1}^{n} d(x)^2 \geq \left(n^{-1/2}\sum_{x=1}^{n} d(x)\right)^2 = n^{-1}\left(\sum_i |A_i|\right)^2 \geq k^2 n^{1-2\delta}.$$

□

Lemma 42. *Let $A \subset \mathbb{R}$, and let L be a set of lines in $\mathbb{R}^2$ such that each line in L is $n^{1-\delta}$-rich in $A \times A$. Then, for at least $\frac{1}{2}|L|^2 n^{-2\delta}$ pairs of lines $(\ell, \ell') \in L \times L$, $\ell^{-1} \circ \ell$ is $\frac{1}{2}n^{1-2\delta}$-rich in $A \times A$.*

Proof. For each line $\ell : y = \lambda x + b$, let $X(\ell) = \{x \in A : \lambda x + b \in A\}$, and similarly let $Y(\ell) = \{y \in A : \lambda^{-1}(y - b) \in A\}$. Observe that $Y(\ell) = X(\ell^{-1})$. Thus, for $(\ell, \ell') \in L \times L$, if $|Y(\ell) \cap Y(\ell')| \geq \frac{1}{2}n^{1-2\delta}$, then $\ell^{-1} \circ \ell'$ is $\frac{1}{2}n^{1-2\delta}$-rich in $A \times A$. Observe that each $Y(\ell)$ has size at least $n^{1-\delta}$. By Lemma 41, at least $\frac{1}{2}|L|^2 n^{-2\delta}$ pairs of sets $(Y(\ell), Y(\ell'))$ have intersection of size at least $\frac{1}{2}n^{1-2\delta}$. □

We define the operation $*$ by $\ell_1 * \ell_2 = \ell_1^{-1} \circ \ell_2$. This formalizes the notion described earlier of combining rich lines in L to form new rich lines (at the cost of a small amount of richness). Given two sets L, L' of $n^{1-\delta}$-rich lines, we would like to consider the set of lines $\ell * \ell'$ which retain a large amount of richness in $A \times A$.

$$\left\{\ell * \ell' : \ell \in L, \ell' \in L', |\ell * \ell' \cap (A \times A)| \geq n^{1-2\delta}/2\right\}. \tag{1}$$

Corollary 43. *Given a set L of lines which are $n^{1-\delta}$-rich in $A \times A$, there exist at least $\frac{1}{2}|L| n^{-2\delta}$ distinct lines of the form $\ell * \ell'$ which are $\frac{1}{2}n^{1-2\delta}$-rich in $A \times A$.*

Proof. There can be at most $|L|$ pairs which map to a given line in $L * L$, or else there exists $\ell_1 \in L$ and $\ell_2 \neq \ell_3$ such that $\ell_1^{-1} \circ \ell_2 = \ell_1^{-1} \circ \ell_3$, a contradiction. By Lemma 42, the result therefore follows. □

In addition to the richness of our new lines in $A \times A$, we will want to have control over the number of pairs (ℓ_1, ℓ_2) which map to the same line under $*$. Let $\mathcal{P}(\ell)$ denote the set of pairs (ℓ_1, ℓ_2) such that $\ell_1 * \ell_2 = \ell$; if X is a set of lines, let $\mathcal{P}(X)$ be the union of the sets $\mathcal{P}(\ell)$ over all $\ell \in X$. For each $0 \leq i \leq \lceil \log_2 |L| \rceil$, let L_i be the set of those lines ℓ in the set (1) such that

$$2^{i-1} < |\mathcal{P}(\ell)| \leq 2^i,$$

and let

$$N_i = \sum_{\ell \in L_i} |\mathcal{P}(\ell)|.$$

Then

$$N_0 + N_1 + \cdots + N_K = \# \left\{ (\ell, \ell') : |\ell * \ell' \cap (A \times A)| \geq \frac{1}{2} n^{1-2\delta} \right\} \geq \frac{1}{2} |L|^2 n^{-2\delta}$$

by Lemma 42. By the pigeonhole principle, at least one N_i satisfies

$$N_i \geq \frac{|L|^2 n^{-2\delta}}{2 \log_2 |L|}.$$

For the maximal such i, we define $L * L$ to be

$$L * L := \{\ell * \ell' : (\ell, \ell') \in \mathcal{P}(L_i)\}.$$

If L is a set of $n^{1-\delta}$ rich lines, we recursively define the sequence of j-fold $*$ operations on L as follows: take $L^{*2} := L * L$ and $L^{*j} := L^{*(j-1)} * L^{*(j-1)}$. We remark that the operation $*$ is not associative: for example, $(L*L)*(L*L)$ will not in general equal $((L * L) * L) * L$.

2.3 Lines in Near-General Position

The following theorem illustrates the behavior of a near-general position set of lines under the operation of $*$.

Theorem 44. *For all $0 < \varepsilon < 1$, there exists $0 < \alpha_0 < \varepsilon$ such that for all $0 < \alpha < \alpha_0$, there exists $0 < \delta_0 < \alpha$ such that for all $0 < \delta < \delta_0$ and for finite sets A with $|A| = n$ sufficiently large, the following holds:*

Let L be a set of at least n^{ε} lines in near-general position (with star families bounded in size by some constant $C \geq 2$ independent of n) which are $n^{1-\delta}$-rich in $A \times A$.

(i) If $L * L$ contains a family P of parallel lines, then $|P| \leq 2 |L * L| n^{2\delta} / |L|$.

(ii) If $L * L$ contains a star family S, then $|S| \leq 2C |L * L| n^{2\delta} / |L|$.

(iii) If P_λ denotes the set of lines in $L * L$ with common slope λ, then $|P_\lambda| \geq n^{\alpha}$ for at most n^{α} numbers λ.

(iv) If S_p denotes the set of lines in $L * L$ with common meeting point p, then $|S_p| \geq n^\alpha$ for at most n^α points p.

Conditions (i), (ii), and (iv) will be shown in this paper. Condition (iii) is shown to hold in [6].

The proofs of conditions (i) and (ii) are similar. Given a line $\ell \in L * L$, recall that $\mathcal{P}(\ell)$ denotes the set of pairs $(\ell_1, \ell_2) \in L \times L$ such that $\ell_1 * \ell_2 = \ell$. If $X \subset L * L$, define $\mathcal{P}(X) := \bigcup_{\ell \in X} \mathcal{P}(\ell)$.

2.3.1 Large Families of Parallel Lines

Proof of Theorem 44(i). Suppose there is a family $P \subseteq L * L$ of parallel lines with

$$|P| > \frac{2\,|L * L|}{|L|} n^{2\delta} \geq \frac{|L|}{2^i},$$

where i is the maximal index between 0 and $\lceil \log_2 |L| \rceil$ such that

$$\#\left\{\ell \in L * L : 2^{i-1} < |\mathcal{P}(\ell)| \leq 2^i\right\} \geq \frac{|L|^2\, n^{-2\delta}}{2 \log_2 |L|}.$$

Then the total number of pairs mapping to lines in P under $*$ will be

$$|\mathcal{P}(P)| = \sum_{\ell \in P} |\mathcal{P}(\ell)| > |P|\, 2^i > |L|\,.$$

Since the lines of L have distinct slopes, it follows that there exist two distinct pairs $(\lambda x + b, \lambda' x + b')$ and $(\lambda x + b, \lambda'' x + b''')$ which each map to some line in P. But then $\lambda' = \lambda''$, and by the distinctness of the pairs it follows that $b' \neq b'''$. Thus, two lines in L are parallel, contradicting the hypothesis that they are in near-general position. □

2.3.2 Large Star Families

Proof of Theorem 44(ii). It is sufficient to consider the case that the lines in S intersect on the y-axis. If not, suppose the center of S is (x_0, y_0), and consider the grid $A' \times A'$, where A' is the translate $A - x_0$. Suppose the pair (ℓ_1, ℓ_2) of $n^{1-\delta}$-rich lines in $A \times A$ maps to a line in S, where $\ell_1 : y = \lambda_1 x + b_1$ and $\ell_2 : y = \lambda_2 x + b_2$. Then

$$\ell_1 * \ell_2 : y = \frac{\lambda_2}{\lambda_1} x + \frac{b_2 - b_1}{\lambda_1}$$

contains the point (x_0, y_0). Let ℓ_1', ℓ_2' be the translates of ℓ_1, ℓ_2 down by x_0 and left by x_0; then

$$\ell_1' : y = \lambda_1 x + \lambda_1 x_0 + b_1 - x_0 \quad \text{and} \quad \ell_2' : y = \lambda_2 x + \lambda_2 x_0 + b_2 - x_0.$$

So (ℓ_1', ℓ_2') maps to

$$\ell_1' * \ell_2' : y = \frac{\lambda_2}{\lambda_1} x + \frac{(\lambda_2 - \lambda_1)x_0 + b_2 - b_1}{\lambda_1}.$$

At $x = 0$, we have

$$y = \frac{\lambda_2}{\lambda_1} x_0 + \frac{b_2 - b_1}{\lambda_1} - x_0 = y_0 - x_0.$$

That is to say, (ℓ_1', ℓ_2') maps to a rich line passing through the point $(0, y_0 - x_0)$. Thus, given a star family of lines $\frac{1}{2}n^{1-2\delta}$-rich in $A \times A$, we can construct a new $\frac{1}{2}n^{1-2\delta}$-rich star family of the same size in the translated grid $A' \times A'$ whose center lies on the y-axis.

Now suppose there is a star family $S \subseteq L * L$ centered at $(0, y_0)$ with

$$|S| > 2C \frac{|L * L|}{|L|} n^{2\delta} \geq C \frac{|L|}{2^i}$$

(where i is taken as in the previous proof). Then the total number of preimages for lines in S will be

$$|\mathcal{P}(S)| = \sum_{\ell \in S} |\mathcal{P}(\ell)| > |S|\, 2^i > C\, |L|\,.$$

Since the lines of L have distinct slopes, it follows that there exist $C+1$ distinct pairs in $L \times L$ mapping to S such that the lines in the first coordinate of the pairs are the same:

$$(\lambda x + b, \lambda_1 x + b_1), (\lambda x + b, \lambda_2 x + b_2), ..., (\lambda x + b, \lambda_{C+1} x + b_{C+1}).$$

Since the y-intercepts of the output lines are the same, it follows that

$$\frac{b_1 - b}{\lambda} = \frac{b_2 - b}{\lambda} = \cdots = \frac{b_{C+1} - b}{\lambda} = y_0;$$

in other words, $b_1 = b_2 = \cdots = b_{C+1} = \lambda y_0 + b$. Thus, L contains $C + 1$ lines with a common y-intercept, contradicting the hypothesis that L is in near-general position with star families bounded in size by C. $\square$

2.3.3 Star Families of Moderate Size

Moving towards a proof of case (iv) of Theorem 44, we begin with a technical sum-product-type result. The result follows from two theorems: a variant of the Balog-Szemerédi-Gowers theorem by Evan Borenstein and Ernie Croot, and the other a result by Croot and Hart on k-fold sumsets when the product set is small.

Theorem 45 ([7])**.** *For every $0 < \varepsilon < 1/2$ and $c > 1$, there exists $\delta > 0$ such that for sufficiently large k and n, the following property holds.*

Let A be a subset of an additive abelian group with $|A| = n$ and $S \subseteq A^k$, the k-fold cartesian product of A with itself. Define

$$\Sigma(S) := \{a_1 + \cdots + a_k : (a_1, ..., a_k) \in S\}.$$

If $|S| \geq |A|^{k-\delta}$ and $|\Sigma(S)| < |A|^c$, then there exists a subset $A' \subseteq A$ with $|A'| \geq |A|^{1-\varepsilon}$ such that for all $h \geq 1$, $|hA'| \leq |A'|^{c(1+h\varepsilon)}$.

Theorem 46 ([26])**.** *For every $h \geq 2$, there exists $\varepsilon_0 = \varepsilon_0(h) > 0$ such that for every $0 < \varepsilon < \varepsilon_0$, there exists $\delta = \delta(\varepsilon) > 0$ such that for sufficiently large n, the following property holds.*

Let $B \subseteq \mathbb{R}$ with $|B| = n$. If $|B.B| \leq |B|^{1+\delta}$, then $|hB| \geq |B|^{\varepsilon}$.

Now let us introduce and prove the lemma:

Lemma 47. *For every $c > 0$ and $k \geq 2$, there exists $\alpha > 0$ such for sufficiently large n, the following property holds.*

If $A_1, A_2, ..., A_k \subseteq \mathbb{R}$ with $|A_i| = n$, $|A_i.A_i| \leq n^{1+\alpha}$ for all $i = 1, ..., k$, and $S \subseteq A_1 \times A_2 \times \cdots \times A_k$ has size $|S| \geq n^{k-\alpha}$, then $|\Sigma(S)| \geq n^c$.

Proof. Let $c > 0$, and let $A_1, ..., A_k, S$ be sets as in the statement of the lemma such that $|\Sigma(S)| < n^c$ (where α, k, and n will be determined later). Let $\varepsilon > 0$, choose

$\delta = \delta(\varepsilon)$ be as in Theorem 45, let $\alpha < \delta/2$, and let $A = A_1 \cup \cdots \cup A_k$. Observe that $n \leq |A| \leq kn$. Note that $S \subset A^k$,

$$|S| \geq n^{k-\alpha} > (|A|/k)^{k-\alpha} > |A|^{k-\delta}, \quad \text{and} \quad |\Sigma(S)| < n^c \leq |A|^c.$$

So there is a subset $A' \subseteq A$ with $|A'| \geq |A|^{1-\varepsilon}$ such that for all $h \geq 1$, $|hA'| < |A|^{c(1+h\varepsilon)}$.

By the pigeonhole principle, A' intersects some A_j, $1 \leq j \leq k$, in a set of size at least $|A'|/k > |A|^{1-\varepsilon}/k$. Let A'' be that intersection, and note that this set satisfies the following:

$$|A''| > |A|^{1-\varepsilon}/k, \quad |A''.A''| \leq |A_j.A_j| \leq n^{1+\alpha}, \quad \text{and} \quad |hA''| < N^{c(1+h\varepsilon)}.$$

Expressing all of this in terms of $|A''|$, we get

$$|A''.A''| \leq n^{1+\alpha} \leq (k|A''|)^{(1+\alpha)/(1-\varepsilon)} \quad \text{and} \quad |hA''| < (k|A''|)^{c(1+h\varepsilon)/(1-\varepsilon)}.$$

Choosing h sufficiently large and α and ε sufficiently small, these inequalities contradict Theorem 46. □

From this lemma we prove a corollary which will have a direct application to the proof of condition (iv) of Theorem 44.

Lemma 48. *There is an absolute constant $c > 0$ such that for every $\varepsilon > 0$, there exists $0 < \alpha_0 < \varepsilon$ such that for all $0 < \alpha < \alpha_0$, there exists $0 < \delta_0 < \alpha$ such that the following holds for all $0 < \delta < \delta_0$ and sufficiently large $n = n(\varepsilon, \alpha, \delta)$:*

If $\{C_1, ..., C_k\}$ is a collection of sets of real numbers such that for all i, $|C_i| \geq n^\alpha$ and $|C_i.C_i| \leq |C_i|^{1+c\delta}$, B is a set of real numbers with $|B| \geq n^\varepsilon$, and $x_1, ..., x_k \in \mathbb{R}$ are distinct constants such that for all i and for each $\lambda \in C_i$, there are at least $|B|^{1-c\delta}$ pairs $(b, b') \in B \times B$ satisfying $\lambda(b - x_i) = b' - x_i$, then $k < n^{\alpha - c\delta}$.

First we need another lemma:

Lemma 49. *Suppose that $d_{i,j}$, $i = 1, ..., k$ and $j = 1, ..., N$ are real numbers satisfying $0 \leq d_{i,j} \leq L$. If $0 \leq C \leq 1$ is defined by*

$$\sum_{i=1}^{k} \sum_{j=1}^{N} d_{i,j} = CLkN,$$

then there exists $i \in \{1, ..., k\}$ such that for at least $\frac{1}{2}kC^2$ indices $i' \in \{1, ..., k\}$,

$$\sum_{j=1}^{N} d_{i,j} d_{i',j} > \frac{1}{2} C^2 L^2 N. \tag{2}$$

Proof. By the Cauchy-Schwarz inequality,

$$\sum_{1 \leq i,i' \leq k} \sum_{j=1}^{N} d_{i,j} d_{i',j} = \sum_{j=1}^{N} \left(\sum_{i=1}^{k} d_{i,j} \right)^2 \geq \frac{1}{N} \left(\sum_{j=1}^{N} \sum_{i=1}^{k} d_{i,j} \right)^2 = C^2 L^2 k^2 N.$$

In particular, there must exist some $i \in \{1, ..., k\}$ such that

$$\sum_{i'=1}^{k} \sum_{j=1}^{N} d_{i,j} d_{i',j} \geq C^2 L^2 k N.$$

Fixing such an i, let T denote the number of indices $i' \in \{1, ..., k\}$ for which

$$\sum_{j=1}^{N} d_{i,j} d_{i',j} \leq \frac{1}{2} C^2 L^2 N.$$

Then

$$\frac{1}{2} T C^2 L^2 N + (k - T) L^2 N \geq C^2 L^2 k N,$$

so

$$T \leq k \frac{1 - C^2}{1 - C^2/2}.$$

Thus, for at least

$$k - T \geq \frac{kC^2}{2 - C^2} \geq \frac{1}{2} kC^2$$

indices $i' \in \{1, ..., k\}$, (2) holds. □

Proof of Lemma 48. By a dyadic pigeonhole argument, there exists a subcollection of the set $\{C_1, ..., C_k\}$ with size $\frac{k}{\log_2(n)}$ and an integer $L \geq n^\alpha$ such that $L \leq |C_i| \leq 2L$

for each C_i in the subcollection. Redefine k to be the number of elements in this subcollection, and reindex so that $C_1, ..., C_k$ are the sets making up the subcollection.

For each $i = 1, ..., k$, construct the directed bipartite graph G_i on vertex set $B_1 \sqcup B_2$ where $B_1 = B_2 = B$ and where (b, b') is an edge if there exists $\lambda \in C_i$ such that $\lambda(b - x_i) = b' - x_i$. Letting $N = |B|$, the sum of the out-degrees in B_1 (and the sum of the in-degrees in B_2) is at least $LN^{1-c\delta}$ by our dyadic pigeonhole argument.

If G is a directed graph, define $\tilde{G}$ to be the graph obtained by reversing the orientation of each of G's edges.

If G and G' are two (2^t+1)-partite directed graphs whose vertex sets are $B_1 \sqcup \cdots \sqcup B_{2^t+1}$, where $B_1 = \cdots = B_{2^t+1} = B$, then define the $(2^{t+1}+1)$-partite directed graph $G \wedge G'$ as follows: Let $V = B_1 \sqcup \cdots \sqcup B_{2^{t+1}+1}$. For $m = 1, ..., 2^t$, let $(b_j, b_{j'}) \in B_m \times B_{m+1}$ be an edge in $G \wedge G'$ if and only if $(b_j, b_{j'}) \in B_m \times B_{m+1}$ is an edge in G. For $m = 2^t + 1, ..., 2^{t+1}$, let $(b_j, b_{j'}) \in B_m \times B_{m+1}$ be an edge in $G \wedge G'$ if and only if $(b_j, b_{j'}) \in B_{m-2^t} \times B_{m+1-2^t}$ is an edge in G'.

Define G_{i_1,i_2} to be $G_{i_1} \wedge \tilde{G}_{i_2}$. By Lemma 49 (taking $C = N^{-c\delta}$) there is an index $1 \leq i_1 \leq k$ such that for at least $\frac{1}{2}kN^{-2c\delta}$ indices $1 \leq i_2 \leq k$,

$$\sum_{j=1}^{N} d_{i_1,j} d_{i_2,j} \geq L^2 N^{1-2c\delta},$$

where $d_{i,j}$ is the number of directed edges in G_i terminating at $b_j \in B_2$. This sum then counts the total number of paths of length 2 in G_{i_1,i_2}, so the average number of paths in G_{i_1,i_2} terminating at a particular $b \in B$ is at least $L^2 N^{-2c\delta}$.

Now, fixing $i_1, ..., i_{t-1}$, we can apply Lemma 49 again to form a (2^t+1)-partite graph

$$G_{i_1,...,i_{t+1}} = G_{i_1,...,i_{t-1},i_t} \wedge \tilde{G}_{i_1,...,i_{t-1},i_{t+1}}$$

with $L^{2^{t-2}} N^{1-O_{t,c}(\delta)}$ length-2^t paths corresponding to 2^t-tuples $(\lambda_1, ..., \lambda_{2^t})$ such that $\lambda_{4m+1} \in C_{i_1}$ for all m.

Using the fact that

$$\lambda_j(\beta_j - x_{i_j}) = \beta_{j+1} - x_{i_j},$$

or, equivalently,

$$\lambda_j\beta_j = \beta_{j+1} - x_{i_j} + \lambda_j x_{i_j},$$

each of these paths corresponds to an equation of the form:

$$\begin{aligned}
\lambda_{2^t}\cdots\lambda_1\beta_1 &= \lambda_{2^t}\cdots\lambda_2(\beta_2 - x_{i_1} + \lambda_1 x_{i_1}) \\
&= \lambda_{2^t}\cdots\lambda_2\beta_2 - \lambda_{2^t}\cdots\lambda_2 x_{i_1} + \lambda_{2^t}\cdots\lambda_1 x_{i_1} \\
&\vdots \\
&= (\beta_{2^t+1} - x_{i_{2^t}}) + \sum_{y=1}^{2^t}\left[\prod_{j=y}^{2^t}\lambda_j\right](x_{i_y} - x_{i_{y-1}}),
\end{aligned}$$

where we define $x_{i_0} = 0$.

By the pigeonhole principle, there exists a choice of β_1 and of the variables λ_j, $j \not\equiv 1 \pmod 4$, for which there are at least $L^{2^{t-2}}N^{-O(\delta)}$ paths in the (2^t+1)-partite graph starting at β_1 and utilizing the edges specified by the selected λ_j (leaving at least $L^{2^{t-2}}$ free choices of edges). Fixing such a β_1 and the variables λ_j except $\lambda_{4s-3} \in C_i$ for all $1 \le s \le 2^{t-2}$, the left-hand side of the equality

$$\lambda_{2^t}\cdots\lambda_1\beta_1 = (\beta_{2^t+1} - x_{i_{2^t}}) + \sum_{y=1}^{2^t}\left[\prod_{j=y}^{2^t}\lambda_j\right](x_{i_y} - x_{i_{y-1}}), \tag{3}$$

is an expression contained in the set

$$C_i^{2^{t-2}} \cdot \beta_1 \cdot \prod_{\substack{1\le j\le 2^t \\ j\not\equiv 1 \pmod 4}} \lambda_j,$$

which by Theorem 2 has size at most $|C_i|^{1+O_t(\delta)}$. Now we rewrite the right-hand side by grouping the terms indexed by y into groups of four, starting at $y = 4r+2$: a typical such group will have the sum

$$\begin{aligned}
\lambda_{4r+5}\lambda_{4r+6}\cdots\lambda_{2^t}\big(\lambda_{4r+2}\lambda_{4r+3}\lambda_{4r+4}(x_{i_{4r+2}} - x_{i_{4r+1}}) + \lambda_{4r+3}\lambda_{4r+4}(x_{i_{4r+3}} - x_{i_{4r+2}}) + \\
\lambda_{4r+4}(x_{i_{4r+4}} - x_{i_{4r+3}}) + x_{i_{4r+5}} - x_{i_{4r+4}}\big).
\end{aligned} \tag{4}$$

Conveniently, all the terms in the parentheses involve products of λ_js, where $j \not\equiv 1 \pmod 4$. Since consecutive pairs of $x_{i_{4r+1}}, x_{i_{4r+2}}, x_{i_{4r+3}}, x_{i_{4r+4}}$ are all distinct (as they are each associated to different star families), $x_{i_{4r+2}} - x_{i_{4r+1}}$, $x_{i_{4r+3}} - x_{i_{4r+2}}$, $x_{i_{4r+4}} - x_{i_{4r+3}}$, and $x_{i_{4r+5}} - x_{i_{4r+4}}$ are all nonzero. It follows that for any choice of two of the parameters among $\lambda_{4r+2}, \lambda_{4r+3}, \lambda_{4r+4}$, there is at most one possible choice of the remaining parameter that can make the expression (4) equal to 0. In fact, the number of paths through the graph resulting in a selection of the λ_js where at least one of the 2^{t-2} four-tuples equals 0 is at most $L^{2^t-1}N^{1-O(\delta)}$. But since there are many more paths than this, there is a choice for the λ_j, $j \not\equiv 1 \pmod 4$, where all the quadruples are nonzero. Re-expressing the right-hand side of (3) in terms of these quadruples for these fixed choices of λ_j, $j \not\equiv 1 \pmod 4$, we find that it is contained in the set

$$\beta_{2^t+1} - x_{i_{2^t}} + \sum_{j=1}^{2^{t-2}} \kappa_j C_i^{4j},$$

where $\kappa_j \neq 0$ are constants. Furthermore, it turns out that for at least $|C_i|^{2^{t-2}-O(\delta)}$ vectors $(c_1, c_2, ..., c_{2^{t-2}}) \in C_i \times C_i^{(2)} \times \cdots \times C_i^{(2^{t-2})}$, this expression is among the expressions in the right-hand side of (3) that we can produce by Theorem 2 (since $|C_i.C_i| \approx |C_i|$). Applying Lemma 47, we find that for t large enough, the number of right-hand side expressions exceeds L^2. This contradicts the fact that the number of left-hand side expressions is bounded by $L^{1+O_{c,t}(\delta)}$. This contradiction finishes the proof. □

Now, we finally establish that if L is in near-general position, then $L * L$ does not contain too many star families of "moderate" size.

Proof of Theorem 44(iv). Suppose that there exist $k = n^\alpha$ star families $S_1, ..., S_k \subseteq L * L$ with $|S_i| \geq n^\alpha$ for all i. Let δ' be an auxiliary parameter such that $\delta < \delta' < \alpha$. We will construct sets $B, C_1, ..., C_k$ and distinct constants $x_1, ..., x_k \in \mathbb{R}$ such that $|C_i| \geq n^\alpha$, such that $|C_i.C_i| \leq |C_i|^{1+O(\delta')}$ and such that for all i and for all $\lambda \in C_i$, we

have $|B|^{1-O(\delta')}$ pairs $(b, b') \in B \times B$ such that $\lambda(b - x_i) = b' - x_i$. This construction is forbidden by Lemma 48, giving us a contradiction.

Begin by taking the x_i to be the x-coordinates of the centers of the star families $S_1, ..., S_k$ in $L * L$. Our first difficulty will be to show that the x_i are distinct. Indeed, this may not be the case, for it is possible that many of the star families lie on common vertical lines. However, suppose that there are K distinct vertical lines on which there are star families. Then there is some such line with at least n^α/K star families on it. Now, since a line $\ell : y = \lambda x + b$ is in $L * L$ if and only if its inverse $\ell^{-1} : x = \lambda y + b$ is also in $L * L$, it follows that there is a horizontal line with n^α/K star families on it, implying there are at least that many distinct vertical lines. Hence, $K \geq n^\alpha/K$, so $K \geq n^{\alpha/2}$. By choosing one star family from each vertical line and ignoring the rest (and reducing α to $\alpha/2$) we attain distinct x-coordinates for the centers of the star families.

Now, fix a star family S_i, and let Λ_i be the set of slopes of the lines in S_i. Observe that lines in $S_i^{*2} := S_i * S_i$ will have slopes in the ratio set $Q_i := \Lambda_i/\Lambda_i$, lines in $S_i^{*3} := S_i^{*2} * S_i^{*2}$ will have slopes in $Q_i^2 = (\Lambda_i/\Lambda_i)^2$, and (in general) lines in $S_i^{*(j+2)}$, $j \geq 0$, will have slopes in the set $Q_i^{2^j}$. Now, not all elements of $Q_i^{2^j}$ will be slopes of lines in $S_i^{*(j+2)}$ (because some combined lines will not be rich enough in $A \times A$). Let $M_{i,j} \subset Q_i^{2^j}$ be the set of slopes of lines in $S_i^{*(j+2)}$.

Observe that $S_i * S_i$ is itself a star family centered at (x_i, x_i). The line $y = x$ is an n-rich line passing through this point, so if $S_i * S_i$ does not contain the line $y = x$, we may add it to $S_i * S_i$ while preserving the fact that all lines in $S_i * S_i$ are rich. Therefore, $S_i^{*j} \subseteq S_i^{*(j+1)}$ for all $j \geq 2$; hence $M_{i,j-1} \subseteq M_{i,j}$ for all $j \geq 1$. Moreover, since $y = x$ is in $S_i * S_i$, $M_{i,j}$ is closed under taking reciprocals for all j. Further note that lines in S_i^{*j} will be $n^{1-2^{O(j)}\delta}$-rich in $A \times A$.

Recall that $\delta < \delta' < \alpha$. Suppose $|M_{i,j+1}| \geq |M_{i,j}|\, n^{\delta'\alpha}$ for all j up to $m = \lfloor 2/\delta'\alpha \rfloor$. Redefining δ if necessary, we may take $1 - 5^m\delta > 0$. For sufficiently large n, we then

have

$$|M_{i,m+1}| \geq n^{\alpha+m\delta'\alpha} \geq n^2.$$

But, since each element of $M_{i,m+1}$ corresponds to a distinct rich line in $A \times A$, this violates Theorem 6. Therefore there exists some $j = j(i) < 2/\delta'\alpha$ such that

$$|M_{i,j+1}| < |M_{i,j}|\, n^{\delta'\alpha} \leq |M_{i,j}|^{1+\delta'}.$$

Therefore, by Lemma 42, there are at least $|M_{i,j}|^{2-O(\delta')}$ pairs $(m, m') \in M_{i,j} \times M_{i,j}$ such that $m/m' \in M_{i,j+1}$. So the multiplicative energy of $M_{i,j}$ satisfies

$$E(M_{i,j}, M_{i,j}) \geq |M_{i,j}|^{3-O(\delta')}.$$

By Theorem 1, we conclude there is a subset $M_i' \subseteq M_{i,j}$ with small multiplicative doubling: $|M_i'.M_i'| \leq |M_i'|^{1+O(\delta')}$. Let S_i' be those lines $\ell \in S_i^{*(j+2)}$ such that the slope of ℓ is in M_i'.

Now, let $\ell : \lambda x + (x_i - \lambda x_i) = \lambda(x - x_i) + x_i$ be a line in the stable star family S_i'. Since this line is rich in $A \times A$, there are $|A|\, n^{-O(\delta')}$ pairs $(a, a') \in A \times A$ such that $\lambda(a - x_i) = a' - x_i$. Taking $C_i = M_i'$ and $B = A$, we obtain the sets forbidden by Lemma 48. □

2.4 Extracting a Near-General Position Set of Lines

Using Theorem 44, we can find a subset of lines in $L * L$ which is in near-general position: the subset will contain no two parallel lines, and all star families in the subset have size bounded by a constant C independent of n.

Corollary 50. *For all $0 < \varepsilon < 1$ there exists $0 < \alpha_0 < \varepsilon$ such that, for all $0 < \alpha < \alpha_0$, there exists $0 < \delta_0 < \alpha$ such that for all $0 < \delta < \delta_0$ and for sufficiently large n, the following holds:*

Let $A \subseteq \mathbb{R}$ be a finite set with $|A| = n$, and let L be a set of at least n^ε lines which are all $n^{1-\delta}$-rich in $A \times A$. If L contains no parallel lines and all star families in L

are bounded above in size by $C = C(\varepsilon, \alpha)$, then there exists a subset $R \subseteq L * L$ such that

- $|R| \geq |L|\, n^{-c\alpha}$ *for some absolute constant c,*
- *R contains no two lines which are parallel, and*
- *at most $k = \lceil \varepsilon/\alpha \rceil$ lines of R pass through any given point of $\mathbb{R}^2$.*

We need a short lemma, which is easily proved by induction.

Lemma 51. *Let k be a nonnegative integer and $0 < \gamma < 1$. Then*

$$\lim_{x\to\infty} x^{k\gamma} \cdot \frac{\binom{x}{x^{1-\gamma}-k}}{\binom{x}{x^{1-\gamma}}} = 1.$$

Proof of Corollary 50. By Theorem 44(iii), there are at most n^α families of parallel lines in $L * L$ with size greater than n^α. By Theorem 44(i), none of these families can have size greater than $2\,|L * L|\, n^{2\delta}/\,|L|$. Thus, deleting all of these lines from $L * L$ leaves us with a set of at least

$$|L * L| - \frac{2\,|L * L|\, n^{\alpha+2\delta}}{|L|} > \frac{1}{2}\,|L * L|$$

lines.

The remaining families of parallel lines in this set have size at most n^α, and these families are all disjoint. By picking a single representative from each family, we form a subset of $L * L$ of at least $\frac{1}{2}\,|L * L|\, n^{-\alpha}$ lines, no two of which are parallel. Invoking Theorem 44(ii) and (iv), we remove from this subset all star families of size greater than n^α to leave us with a subset L' with at least $\frac{1}{4}\,|L * L|\, n^{-\alpha}$ lines.

By Corollary 43, there are at least $|L|\, n^{-3\delta}$ lines in $L * L$, so L' contains at least $|L|\, n^{-2\alpha}$ lines.

Uniformly at random choose a subset $R \subseteq L'$ of $\lceil |L'|\, n^{-c\alpha} \rceil$ lines, where $c > 0$ is a parameter to be chosen later. The probability that a star family S in L' contains at

least k lines of R is

$$\frac{\binom{|S|}{k}\cdot\binom{|L'|-k}{|R|-k}}{\binom{|L'|}{|R|}} \leq \frac{n^{k\alpha}}{k!}\cdot\frac{\binom{|L'|}{|R|-k}}{\binom{|L'|}{|R|}}.$$

Applying Lemma 51 with $x = |L'|$ and $x^{1-\gamma} = |R| = x^{1-c\alpha\log_x(n)}$, for large values of n we have

$$\frac{\binom{|L'|}{|R|-k}}{\binom{|L'|}{|R|}} = (1+o(1))\cdot|L'|^{-kc\alpha\log_{|L'|}(n)} \leq 2n^{-kc\alpha}.$$

Since there are at most $n^{2\varepsilon}$ star families, the expected number of star families with at least k lines of R is bounded by

$$\frac{2}{k!}n^{2\varepsilon+k(1-c)\alpha}.$$

Taking $k = \left\lceil\frac{\varepsilon}{\alpha}\right\rceil$ and $c = 3$ makes this expected value less than 1, meaning there is some choice of R such that no star family has more than $\lceil\varepsilon/\alpha\rceil$ lines in R.

Thus, R is a near-general position subset of L' (and therefore of L) with size at least $|L'|\,n^{-3\alpha} \geq |L|\,n^{-5\alpha}$. □

We remark that the proof still holds if $L * L$ above is replaced by $L'' \subseteq L * L$ so long as $|L''| \geq |L|\,n^{-c_0\alpha}$ for some $c_0 > 0$. We will use this modified version in the proof that Theorem 40 implies Theorem 39.

2.5 Proof of the Weakened Theorem

Using Corollary 50, we are now ready to prove Theorem 40. A major tool used will be the commutator graph, which we draw from [30].

Let $A \subseteq \mathbb{R}$, let $\delta > 0$, and let L be a set of $n^{1-\delta}$-rich lines in $\mathbb{R}^2$. The *commutator graph* on L is the graph $G = (V, E)$, where

$$V(G) = L * L \cup L^{-1} * L^{-1}$$

(with the minor change that we require minimum richness only $n^{1-5\delta}$ for each line in $L * L$ and $L^{-1} * L^{-1}$) and

$$E(G) = \left\{\{f * g, g^{-1} * f^{-1}\} : f, g \in L, f * g \in L * L, g^{-1} * f^{-1} \in L^{-1} * L^{-1}\right\}.$$

We draw attention to the fact that the lines $f * g$ and $g^{-1} * f^{-1}$ have the same slope. Hence, any edge of the commutator graph is between two parallel lines.

Proof of Theorem 40. Let $\varepsilon > 0$, let $\delta > 0$ be much smaller than ε, and let $A \subset \mathbb{R}$ with $n = |A| > 0$. Suppose for a contradiction that L is a set of at least $n^{1-\varepsilon}$ lines, all $n^{1-\delta}$-rich in $A \times A$, and that L is in near-general position with star families bounded in size by a constant $C > 0$ independent of n. Consider the commutator graph on L.

If $|V(G)| \geq n^{1+4\delta}$, then we contradict Theorem 6, so let us assume that $|V(G)| < n^{1+4\delta}$. We claim that $|E(G)| \geq n^{2-6\delta}$. If this is true, then there is a vertex with degree at least $|E(G)| / |V(G)|$, so there is a connected component (corresponding to a set of parallel lines) of size $n^{1-10\delta}$, in contradiction with Theorem 44(i).

Let $S(f) = X(f) \times Y(f)$ for each $f \in L$. By applying Lemma 41 to the collection of sets $S(f)$, where each set $S(f)$ has size at least $n^{2-2\delta}$, we must have at least $n^{2-4\delta}/2 \geq n^{2-5\delta}$ pairs $S(f), S(g)$ with $|S(f) \cap S(g)| \geq n^{2-4\delta}/2 \geq n^{2-5\delta}$. Note that for any sets A_1, A_2, A_3, A_4, $(A_1 \times A_3) \cap (A_2 \times A_4) = (A_1 \cap A_2) \times (A_3 \cap A_4)$. Thus, since $|S(f) \cap S(g)| \geq n^{2-5\delta}$, we have $|X(f) \cap X(g)| \geq n^{1-5\delta}$ and $|Y(f) \cap Y(g)| \geq n^{1-5\delta}$: that is, there are at least $n^{2-5\delta}$ pairs $f, g \in L$ such that $f * g$ and $g^{-1} * f^{-1}$ are each $n^{1-5\delta}$-rich.

Let f_i, g_i denote the lines such that $P_i := \{f_i * g_i, g_i^{-1} * f_i^{-1}\}$ is a pair of $n^{1-5\delta}$-rich lines. Given an index i, f_i and g_i intersect at a unique point (x, y); it then follows that x is the unique fixed point of $f_i * g_i$ and y is the unique fixed point of $g_i^{-1} * f_i^{-1}$. Suppose there were $2C + 2$ indices $i_1, ..., i_{2C+2}$ such that $P_{i_j} = P_{i_k}$ for all $1 \leq j, k \leq 2C + 2$. Then there would exist $C + 1$ indices $i_{j_1}, ..., i_{j_{C+1}}$ such that

$$f_{i_{j_1}} * g_{i_{j_1}} = \cdots = f_{i_{C+1}} * g_{i_{C+1}} \quad \text{and} \quad g_{i_{j_1}}^{-1} * f_{i_{j_1}}^{-1} = \cdots = g_{i_{C+1}}^{-1} * f_{i_{C+1}}^{-1}.$$

Since for each $1 \leq k \leq C + 1$ there is a unique (x, y) such that $f_{i_{j_k}} * g_{i_{j_k}}(x) = x$ and $g_{i_{j_k}}^{-1} * f_{i_{j_k}}^{-1}(y) = y$, it follows that $f_{i_{j_k}}$ and $g_{i_{j_k}}$ all intersect the point (x, y). Since the $f_{i_{j_k}} * g_{i_{j_k}}$ must all have the same slope and L has no parallel lines, we cannot have

that $f_{i_{j_k}} = f_{i_{j'_k}}$ for $k \neq k'$ or else $g_{i_{j_k}} = g_{i_{j'_k}}$ as well, contradicting distinctness of the pairs. Similarly we must have $g_{i_{j_k}} \neq g_{i_{j'_k}}$ for $k \neq k'$. The collection

$$\{f_{i_{j_k}} : 1 \leq k \leq C+1\} \cup \{g_{i_{j_k}} : 1 \leq k \leq C+1\}$$

must therefore contain at least $C+1$ distinct lines (a single line may appear as an f_{i_j} at most once and as a g_{i_j} at most once). But then we have a set of more than C concurrent lines at (x, y), contradicting the hypothesis that L is in almost-general position.

Thus, for each edge e, there are at most $2C+2$ pairs $\{f_i \circ g_i^{-1}, g_i^{-1} \circ f_i\}$ equal to e, so the total number of edges in G is at least $n^{2-5\delta}/(2C+2) \gg n^{2-6\delta}$, yielding a contradiction with Theorem 44(i). □

We remark that there is a constant $0 < c < 1$ such that taking $\delta = c\varepsilon$ is sufficient for the proof to go through.

2.6 Proof of the Main Theorem

For $\ell \in L * L$, recall that $\mathcal{P}(\ell)$ is the set of all pairs $(f, g) \in L \times L$ such that $f * g = \ell$.

Lemma 52. *For all $0 < \varepsilon < 1$, there exists $0 < \alpha_0 < \varepsilon$ such that, for all $0 < \alpha < \alpha_0$, there exists $0 < \delta_0 < \alpha$ such that for all $0 < \delta < \delta_0$ and for sufficiently large n, the following holds:*

Let $A \subseteq \mathbb{R}$ have size n, and let L be a set of at least n^ε near-general position lines, all of which are $n^{1-\delta}$-rich in $A \times A$. Then there exists a set $L' \subseteq L * L$ such that L' is a set of lines in near-general position, $|L'| > |L|\, n^{-5\alpha-4\delta}$, and for all $\ell \in L'$,

$$|\mathcal{P}(\ell)| \geq \frac{|L|^2\, n^{-3\delta}}{2\,|L * L|}.$$

Proof. Let

$$S := \left\{(f, g) \in L \times L : f * g \text{ is } \frac{1}{2}n^{1-4\delta}\text{-rich}\right\} \cap \mathcal{P}(L_i),$$

where $\mathcal{P}(L_i)$ is as in our definition of $L * L$. Then $|S| \geq \frac{|L|^2 n^{-2\delta}}{2\log_2|L|} \gg |L|^2 n^{-3\delta}$. Let

$$T := \left\{ (f,g) \in S : |\mathcal{P}(f*g)| \leq \frac{|L|^2 n^{-3\delta}}{2|L*L|} \right\}.$$

If $|T| > \frac{1}{2}|S|$, then we obtain an absurdity:

$$|L*L| = \sum_{(f,g)\in S} \frac{1}{|\mathcal{P}(f*g)|} = \sum_{(f,g)\in S\setminus T} \frac{1}{|\mathcal{P}(f*g)|} + \sum_{(f,g)\in T} \frac{1}{|\mathcal{P}(f*g)|} \geq$$
$$\frac{1}{|L|}|S\setminus T| + \frac{2|L*L|n^{3\delta}}{|L|^2}|T| > |L*L|.$$

Thus, $|S\setminus T| \geq |L|^2 n^{-3\delta}/2 > |L|^2 n^{-4\delta}$. Letting $L' = \{f*g : (f,g) \in S\setminus T\}$, we then have $|L'| \geq |L| n^{-4\delta}$. Apply Corollary 50 to deduce that L' contains a subset of $|L| n^{-5\alpha-4\delta}$ lines in near-general position. □

Proof of Theorem 39. Let L be a set of n^ε lines in general position, all of which are $n^{1-\delta}$-rich for some $\delta > 0$ to be chosen later. Fix $\alpha < \varepsilon$, and suppose $\left|L^{*(k+1)}\right| \geq \left|L^{*k}\right| n^{5\alpha}$ for all k up to $m = \lfloor 2/\alpha \rfloor$. (By Corollary 50, we may further assume that L^{*j} is in near-general position for all $j \leq k$ at the cost of a factor of $n^{4\alpha}$ each iteration.) Redefining δ if necessary, we can take $1 - 4\cdot 5^m\delta > 0$. For sufficiently large n, we then have

$$\left|L^{*(m+1)}\right| \geq n^{\varepsilon+m\alpha} \geq n^2.$$

But this violates Theorem 6, so such an m cannot exist. Therefore there exists $k < 2/\alpha$ such that

$$\left|L^{*(k+1)}\right| < \left|L^{*k}\right| n^{5\alpha}.$$

In this case, let $L' = L^{*k}$ for the smallest such k (such that the above inequality would now read $|L'*L'| < |L'| n^{5\alpha}$), let $\alpha' < 5\alpha$ such that $\alpha' \ll \varepsilon$, let $N = |L'|$, and choose $\delta' \leq 5^k\delta$ such that $\delta' \ll \alpha'$.

By applying Lemma 52, we can restrict our attention to a subset $L'' \subseteq L'*L'$ of size at least $Nn^{-5\alpha'-4\delta'}$ such that all lines in L'' are in near-general position and, for

all $\ell \in L''$,

$$|\mathcal{P}(\ell)| \geq \frac{N^2}{2\,|L' * L'|\, n^{3\delta'}} \geq \frac{N}{2n^{5\alpha+3\delta'}} > \frac{N}{2n^{\alpha'+3\delta}}.$$

If ℓ is a line in L'', then $\ell = f * g$ for some $f, g \in L'$. We will then have at least

$$\frac{1}{C}\,|L''|\,(Nn^{-\alpha'-4\delta'})^2 \geq \frac{1}{C}N^3 n^{-5\alpha'-8\delta'} \gg N^3 n^{-6\alpha'}$$

solutions $(f, g, f', g') \in L \times L \times L \times L$ to the equation

$$f' * f(0) = g' * g(0) \tag{5}$$

(The factor of $1/C$ comes from the fact that L'' is a set of lines in near-general position, so at most C lines will share a y-intercept.)

Now, fixing f', g' in (5) and letting f, g vary, we can interpret (5) as the line $f' * g'$, where the x and y variables are the y-intercepts of f and g. Letting B be the set of y-intercepts among lines in L'', we may interpret the above count of solutions to (5) as stating that many of the lines $f' * g'$ are $Nn^{-7\alpha'}$-rich in the new grid $B \times B$. Indeed, let

$$S := \{(f', g') \in L' \times L' : f' * g' \text{ is } Nn^{-7\alpha'}\text{-rich in } B \times B\},$$

and let $p(f', g')$ denote the number of points that $f' * g'$ intersects in $B \times B$. Then, for a contradiction, assume $|S| < N^2 n^{-8\alpha'}$. This implies the absurdity:

$$N^3 n^{-6\alpha'} = \sum_{(f',g')\in S} p(f', g') + \sum_{(f',g')\in S^c} p(f', g') < |S|\,|B| + |S^c|\,(Nn^{-7\alpha'}) <$$
$$|S|\,Nn^{\alpha'} + (N^2 - |S|)Nn^{-7\alpha'} < |S|\,Nn^{\alpha'} + N^3 n^{-7\alpha'} < 2N^3 n^{-7\alpha'} \ll N^3 n^{-6\alpha'}.$$

(Note that this requires $N \gg n^{4\alpha'}$, which is satisfied provided $\alpha' \ll \varepsilon$ because $N \gg n^{\varepsilon}$.) Thus, $|S| \geq N^2 n^{-8\alpha'}$, and that implies that we have at least $Nn^{-8\alpha'}$ lines that are all $Nn^{-7\alpha'}$-rich in $B \times B$. Moreover, since L' is in near-general position, we may extract a set L''' from $\{f' * g' : (f', g') \in S\} \subseteq L' * L'$ that is in near-general position, and L''' has size at least $Nn^{-11\alpha'}$. However, this is in contradiction with Theorem 40,

since L''' is a set of $N^{1-\gamma}$-rich lines in near-general position for some $\gamma > 0$, and $|L'''| \geq N^{1-O(\gamma)}$. $\square$

CHAPTER III

FEW PRODUCTS, MANY DIFFERENCES

In this chapter, we give a short proof of the following conditional result:

Theorem 53. *If Conjecture 37 holds, then there exists an absolute constant $c > 0$ such that, for every $0 < \varepsilon < \frac{1}{2}$, there exists $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that, for all $A \subset \mathbb{R}$ with $|A| \geq n_0$ and $|A.A| \leq |A|^{1+\varepsilon}$,*

$$|A - A| \geq c\frac{|A|^{2-3\varepsilon}}{(\log |A|)^8}.$$

A straightforward corollary of Conjecture 37 that we will use in the proof of Theorem 53 is:

Corollary 54. *If $A \subset \mathbb{R}$, then*

$$\#\{(x_1, ..., x_8) \in A^8 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\} = O(|A|^6 \log |A|).$$

Proof. Letting

$$r(s) = \#\{(\mathbf{v}, \mathbf{w}) \in A^2 \times A^2 : |\mathbf{v} \times \mathbf{w}| = s\},$$

we see that

$$r(s)^2 = \#\{(x_1, ..., x_8) \in A^8 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8 = s\}.$$

There are at most $|A|^3$ linearly dependent pairs of vectors in $A \times A$ ($|A|^2$ choices for the first vector, and then $|A|$ choices of vectors along the line spanned by the first vector). Two vectors $(a, b), (c, d)$ are linearly dependent if and only if $(a, b) \times (c, d) = ad - bc = 0$. In other words, $r(0) \leq |A|^3$, so $r(0)^2 \leq |A|^6$.

We conclude that

$$\#\{(x_1,...,x_8) \in A^8 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\} = \sum_{s>0} r(s)^2 + r(0)^2 =$$
$$O(|A|^6 \log |A|) + |A|^6 = O(|A|^6 \log |A|).$$

□

Proof of Theorem 53. Fix $\varepsilon > 0$ and let $A \subset \mathbb{R}$ be a finite set such that $|A.A| \leq |A|^{1+\varepsilon}$. Let

$$r(p) = \#\{(a,b) \in A \times A : ab = p\}.$$

Then

$$\sum_{p \in A.A} r(p) = |A|^2$$

and

$$\sum_{p \in A.A} r(p)^2 = E_\times(A) = \#\{(a,b,c,d) \in A^4 : ab = cd\},$$

where $E_\times(A)$ is the energy of A as a multiplicative set.

By a dyadic pigeonhole principle argument, there exists $t \in \{0,...,2\log|A|\}$ such that

$$\sum_{\substack{p \in A.A \\ 2^t \leq r(p) < 2^{t+1}}} r(p) \geq \frac{|A|^2}{2\log|A|}.$$

Let $F = \{(a,b) \in A \times A : 2^t \leq r(ab) < 2^{t+1}\}$ and $P = \{ab : (a,b) \in F\} \subseteq A.A$ (the mnemonic being F for "factors" and P for "products"). Then

$$|F| = \sum_{\substack{p \in A.A \\ 2^t \leq r(p) < 2^{t+1}}} r(p) \geq \frac{|A|^2}{2\log|A|}$$

and

$$|F| = \sum_{p \in P_t} r(p) \leq 2^{t+1} |P|.$$

This gives us a lower bound on $2^t |P|$:

$$2^t |P| \geq \frac{|A|^2}{4\log|A|}. \tag{6}$$

Since $P \subseteq A.A$ and $|A.A| \leq |A|^{1+\varepsilon}$, we in turn obtain a lower bound on 2^t:

$$2^t \geq \frac{|A|^{1-\varepsilon}}{4 \log |A|}. \tag{7}$$

Now, for $w \in A$, let $F_w = \{(a,b) \in F : ab \in w.A\}$; that is, F_w is the set of those pairs of factors in F whose product lives in the multiplicative translation of A by w. Then

$$\sum_{w \in A} |F_w| = \sum_{w \in A} \sum_{p \in P} \#\{(a,b) \in F : ab = p\} \mathbf{1}_{w.A}(p).$$

Switching the order of summation, we obtain

$$\sum_{p \in P} \#\{(a,b) \in F : ab = p\} \sum_{w \in A} \mathbf{1}_{w.A}(p) = \sum_{p \in P} \#\{(x,y) \in F : xy = p\}^2 \geq 2^{2t+2} |P|.$$

Thus,

$$\sum_{w \in A} |F_w| \geq 2^{2t+2} |P|. \tag{8}$$

For $w \in A$, define the set O_w by

$$O_w = \{(x_1, ..., x_8) \in F \times F \times F \times F : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8 \text{ and } x_1x_2 \in w.A\}$$

(the mnemonic here is O for "octuples"). First observe that

$$\sum_{w \in A} |O_w| = \sum_{w \in A} \sum_{(x_1,x_2) \in F} \#\{(x_1, ..., x_8) \in P^4 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\} \mathbf{1}_{w.A}(x_1x_2) =$$

—after again changing the order of summation—

$$= \sum_{(x_1,x_2) \in F} \#\{(x_1, ..., x_8) \in P^4 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\} \sum_{w \in A} \mathbf{1}_{w.A}(x_1x_2)$$

$$= \sum_{(x_1,x_2) \in F} \#\{(x_1, ..., x_8) \in P^4 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\} r(x_1x_2)$$

$$\leq 2^{t+1} \#\{(x_1, ..., x_8) \in P^4 : x_1x_2 - x_3x_4 = x_5x_6 - x_7x_8\}$$

$$\leq c2^t |A|^6 \log |A|$$

for some absolute constant $c > 0$ by Corollary 54. Thus,

$$\sum_{w \in A} |O_w| \leq c2^t |A|^6 \log |A|. \tag{9}$$

We now prove a lemma from which Theorem 53 will quickly follow.

Lemma 55. *There is some $w \in A$ and some absolute constant $c' > 0$ such that for sufficiently large $|A|$,*

$$|F_w|^4 > c' \frac{|A|^{2-3\varepsilon}}{(\log |A|)^8} |O_w| .$$

Proof. For a contradiction, suppose that for all $w \in A$,

$$|F_w|^4 \leq c' \frac{|A|^{2-3\varepsilon}}{(\log |A|)^8} |O_w| ,$$

where $c' > 0$ is a parameter to be determined later. Then

$$\sum_{w \in A} |F_w|^4 \leq c' \frac{|A|^{2-3\varepsilon}}{(\log |A|)^8} \sum_{w \in A} |O_w| . \tag{10}$$

Using Hölder's inequality and eq. (8), we obtain

$$\sum_{w \in A} |F_w|^4 \geq \frac{1}{|A|^3} \left(\sum_{w \in A} |F_w| \right)^4 \geq \frac{2^{8t+8} |P|^4}{|A|^3}.$$

In the other direction, using eq. (9) and eq. (10), we obtain

$$\sum_{w \in A} |F_w|^4 \leq c' \frac{|A|^{2-3\varepsilon}}{(\log |A|)^8} \sum_{w \in A} |O_w| \leq cc' \frac{A^{8-3\varepsilon} 2^t}{(\log |A|)^7}$$

for an appropriate constant $c > 0$. Hence

$$cc' \frac{|A|^{11-3\varepsilon}}{(\log |A|)^7} \geq 2^{7t+8} |P|^4 .$$

Using the estimates from eq. (6) and eq. (7) we obtain

$$cc' \frac{|A|^{11-3\varepsilon}}{(\log |A|)^7} \geq \frac{2^8 |A|^{11-3\varepsilon}}{(4 \log |A|)^7},$$

and taking $c' < 1/64c$ yields a contradiction. □

Let w be the element of A and c' the constant obtained from Lemma 55. Then let

$$G = F_w \times F_w = \{(a, b, c, d) \in F \times F : ab \in w.A \text{ and } cd \in w.A\},$$

let

$$n(s) = \#\{(a, b, c, d) \in G : ab - cd = s\},$$

and let

$$S = \{s : n(s) > 0\}.$$

By Cauchy-Schwarz,

$$\sum_{s \in S} n(s)^2 \geq \frac{|G|^2}{|S|}.$$

On the other hand,

$$\sum_{s \in S} n(s)^2 \leq |O_w| \leq c' |F_w|^4 \frac{(\log |A|)^8}{|A|^{2-3\varepsilon}}.$$

Since $|G| = |F_w|^2$, combining the two inequalities yields

$$|S| = \Omega\left(\frac{|A|^{2-3\varepsilon}}{(\log |A|)^8}\right).$$

Finally, observe that $|S| \leq |A - A|$ since for each $s \in S$, we have $s = wa - wb = w(a - b) \in w.(A - A)$.

□

REFERENCES

[1] Alon, N. and Spencer, J. H., *The Probabilistic Method*. Wiley-Interscience, 3rd ed., 2008.

[2] Amoroso, F. and Viada, E., "Small Points on Subvarieties of a Torus," *Duke Mathematics Journal*, vol. 150, pp. 407–442, 2009.

[3] Babai, L., Nikolov, N., and Pyber, L., "Product Growth and Mixing in Finite Groups," in *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 248–257, Society for Industrial and Applied Mathematics, 2008.

[4] Babai, L. and Seress, Á., "On the Diameter of Permutation Groups," *European Journal of Combinatorics*, vol. 13, pp. 231–243, 1992.

[5] Balog, A. and Szemerédi, E., "A Statistical Theorem of Set Addition," *Combinatorica*, vol. 14, pp. 263–268, 1994.

[6] Borenstein, E. and Croot, E., "On Rich Lines in Grids," *Discrete and Computational Geometry*, vol. 43, pp. 824–840, 2010.

[7] Borenstein, E. and Croot, E., "On a Certain Generalization of the Balog-Szeméredi Theorem," *SIAM Journal of Discrete Mathematics*, vol. 25, pp. 685–694, 2011.

[8] Bourgain, J., "On the Erdős-Volkmann and Katz-Tao Ring Conjectures," *Geometric and Functional Analysis*, vol. 13, pp. 334–365, 2003.

[9] Bourgain, J., "More on the Sum-Product Phenomenon in Prime Fields and Its Applications," *International Journal of Number Theory*, vol. 1, pp. 1–32, 2005.

[10] Bourgain, J., "Sum-Product Theorems and Exponential Sum Bounds in Residue Classes for General Modulus," *Les Comptes Rendus Mathématique*, vol. 344, pp. 349–352, 2007.

[11] Bourgain, J. and Chang, M.-C., "On Multiple Sum and Product Sets of Finite Sets of Integers," *Les Comptes Rendus Mathématique*, vol. 337, pp. 499–503, 2003.

[12] Bourgain, J. and Chang, M.-C., "A Gauss Sum Estimate in Arbitrary Finite Fields," *Les Comptes Rendus Mathématique*, vol. 342, pp. 643–646, 2006.

[13] Bourgain, J. and Gamburd, A., "Uniform Expansion Bounds for Cayley Graphs of $\mathrm{SL}_2(\mathbb{F}_p)$," *Annals of Mathematics*, vol. 167, pp. 625–642, 2008.

[14] Bourgain, J. and Garaev, M. Z., "On a Variant of Sum-Product Estimates and Explicit Exponential Sum Bounds in Prime Fields," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 146, pp. 1–21, 2009.

[15] Bourgain, J., Glibichuk, A., and Konyagin, S., "Estimates for the Number of Sums and Products and for Exponential Sums over Subgroups in Fields of Prime Order," *Journal of the London Mathematical Society*, vol. 73, pp. 380–398, 2006.

[16] Bourgain, J., Katz, N. H., and Tao, T., "A Sum-Product Estimate in Finite Fields, and Applications," *Geometric and Functional Analysis*, vol. 14, pp. 27–57, 2004.

[17] Bourgain, J. and Konyagin, S., "Estimates for the Number of Sums and Products and for Exponential Sums over Subgroups in Fields of Prime Order," *Les Comptes Rendus Mathématique*, vol. 337, pp. 75–80, 2003.

[18] Breuillard, E., Green, B., and Tao, T., "Approximate Subgroups of Linear Groups," *Geometric and Functional Analysis*, vol. 21, pp. 774–819, 2011.

[19] Bukh, B. and Tsimerman, J., "Sum-Product Estimates for Rational Functions," *Proceedings of the London Mathematical Society*, vol. 104, pp. 1–26, 2011.

[20] Chang, M.-C., "The Erdős-Szemerédi Problem on Sum Set and Product Set," *Annals of Mathematics*, vol. 157, pp. 939–957, 2003.

[21] Chang, M.-C., "A Sum-Product Theorem in Semi-Simple Commutative Banach Algebras," *Journal of Functional Analysis*, vol. 212, pp. 399–430, 2004.

[22] Chang, M.-C., "Sum and Product of Different Sets," *Contributions to Discrete Mathematics*, vol. 1, 2006.

[23] Chang, M.-C., "Additive and Multiplicative Structure in Matrix Spaces," *Combinatorics, Probability and Computing*, vol. 16, pp. 219–238, 2007.

[24] Chang, M.-C., "Product Theorems in SL_2 and SL_3," *Journal of the Institute of Mathematics of Jussieu*, vol. 7, pp. 1–25, 2008.

[25] Chang, M.-C. and Solymosi, J., "Sum-Product Theorems and Incidence Geometry," *Journal of the European Mathematical Society*, vol. 9, pp. 545–560, 2007.

[26] Croot, E. and Hart, D., "h-fold Sums from a Set with Few Products," *SIAM Journal of Discrete Mathematics*, vol. 24, pp. 505–519, 2010.

[27] Elekes, G., "On Linear Cominatorics I," *Combinatorica*, vol. 17, pp. 447–458, 1997.

[28] Elekes, G., "On the Number of Sums and Products," *Acta Arithmetica*, vol. 81, pp. 365–367, 1997.

[29] Elekes, G., "On Linear Cominatorics II," *Combinatorica*, vol. 18, pp. 13–25, 1998.

[30] Elekes, G., "Sums Versus Products in Number Theory, Algebra and Erdős Geometry," in *Paul Erdős and His Mathematics, II* (Halasz, G., Lovasz, L., Simonovits, M., and Sós, V. T., eds.), vol. 11 of *Bolyai Society Mathematical Studies*, pp. 241–290, Janos Bolyai Mathematical Society and Springer Science+Business Media, 2002.

[31] Elekes, G., Nathanson, M., and Ruzsa, I., "Convexity and Sumsets," *Journal of Number Theory*, vol. 83, pp. 194–201, 2000.

[32] Elekes, G. and Ruzsa, I., "Few Sums, Many Products," *Studia Scientiarum Mathematicarum Hungarica*, pp. 301–308, 1983.

[33] Elekes, G. and Ruzsa, I., "The Structure of Sets with Few Sums Along a Graph," *Journal of Combinatorial Theory, Series A*, vol. 113, pp. 1476–1500, 2006.

[34] Erdős, P. and Szemerédi, E., "On Sums and Products of Integers," *Studies in Pure Mathematics*, pp. 213–218, 1983.

[35] Evertse, J.-H., Schlickewei, H. P., and Schmidt, W. M., "Linear Equations in Variables Which Lie in a Multiplicative Group," *Annals of Mathematics*, vol. 155, pp. 807–836, 2002.

[36] Ford, K., "Sums and Products from a Finite Set of Real Numbers," *Ramanujan Journal*, vol. 2, pp. 59–66, 1998.

[37] Ford, K., "Integers with a Divisor in $[y, 2y)$," in *Anatomy of Integers* (de Koninck, J.-M., Granville, A., and Luca, F., eds.), vol. 46 of *CRM Proceedings Lecture Notes*, pp. 65–80, American Mathematical Society, 2008.

[38] Freiman, G. A., "Addition of Finite Sets," *Doklady Akademii Nauk SSSR*, vol. 158, pp. 1038–1041, 1964.

[39] Freiman, G. A., *Foundations of a Structural Theory of Set Addition*. Kazan. Gosudarstv. Ped. Inst; Elabuzh. Gosudarstv. Ped. Inst., 1966.

[40] Garaev, M. Z., "An Explicit Sum-Product Estimate in $\mathbb{F}_p$," *International Mathematical Research Notices*, vol. 35, pp. 1–11, 2007.

[41] Garaev, M. Z., "The Sum-Product Estimate for Large Subsets of Prime Fields," *Proceedings of the American Mathematical Society*, vol. 136, pp. 2735–2739, 2008.

[42] Gowers, W. T., "Quasirandom Groups," *Combinatorics, Probability and Computing*, vol. 17, pp. 363–387, 2008.

[43] Gowers, W. T., "A New Proof of Szenerédi's Theorem for Arithmetic Progressions of Length Four," *Geometric and Functional Analysis*, vol. 8, 1998.

[44] Green, B., "Sum-Product Phenomena in $\mathbb{F}_p$: A Brief Introduction." arXiv:0904.2075, 2009.

[45] Guth, L. and Katz, N. H., "On the Erdős Distinct Distance Problem in the Plane," *Annals of Mathematics*. To appear.

[46] Hart, D., Iosevich, A., and Solymosi, J., "Sum-Product Estimates in Finite Fields via Kloosterman Sums," *International Mathematics Research Notices*, vol. 2007, 2007.

[47] Helfgott, H. A., "Growth and Generation in $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$," *Annals of Mathematics*, vol. 167, pp. 601–623, 2008.

[48] Helfgott, H. A., "Growth in $\mathrm{SL}_3(\mathbb{Z}/p\mathbb{Z})$," *Journal of the European Mathematical Society*, vol. 13, pp. 761–851, 2011.

[49] Hoory, S., Linial, N., and Wigderson, A., "Expander Graphs and Their Properties," *Bulletin of the American Mathematical Society*, vol. 43, pp. 439–561, 2006.

[50] Iosevich, A., Roche-Newton, O., and Rudnev, M., "On an Application of Guth-Katz Theorem," *Mathematical Research Letters*, vol. 18, pp. 691–697, 2011.

[51] Katz, N. H. and Shen, C.-Y., "Garaev's Inequality in Finite Fields Not of Prime Order," *Online Journal of Analytic Combinatorics*, vol. 3, 2008.

[52] Katz, N. H. and Shen, C.-Y., "A Slight Improvement to Garaev's Sum Product Estimate," *Proceedings of the American Mathematical Society*, vol. 136, pp. 2499–2504, 2008.

[53] Konyagin, S. and Rudnev, M., "New Sum Product Type Estimates." arXiv:1207.6785, 2013.

[54] Nathanson, M., "On Sums and Products of Integers," *Proceedings of the American Mathematical Society*, vol. 125, pp. 9–16, 1997.

[55] Nikolav, N. and Pyber, L., "Product Decompositions of Quasirandom Groups and a Jordan-Type Theorem," *Journal of the European Mathematical Society*, vol. 13, pp. 1063–1077, 2011.

[56] Petridis, G., "New Proofs of Plünnecke-type Estimates for Product Sets in Groups," *Combinatorica*, vol. 32, pp. 721–733, 2012.

[57] Plünnecke, H., "Eine zahlentheoretische Anwendung der Graphtheorie," *Journal für die reine und angewandte Mathematik*, vol. 243, pp. 171–183, 1970.

[58] Ruzsa, I., "An Application of Graph Theory to Additive Number Theory," *Scientia, Series A*, vol. 3, pp. 97–109, 1989.

[59] Ruzsa, I., "Addendum to: An Application of Graph Theory to Additive Number Theory," *Scientia, Series A*, vol. 4, pp. 93–94, 1990/1991.

[60] Ruzsa, I., "Generalized Arithmetical Progressions and Sumsets," *Acta Mathematica Hungarica*, vol. 65, pp. 379–388, 1994.

[61] Schwartz, R. and Solymosi, J., "Combinatorial Applications of the Subspace Theorem." arXiv:1311.3743, 2013.

[62] Solymosi, J., "On Sum-Sets and Product-Sets of Complex Numbers," *Journal de Théorie des Nombres de Bordeaux*, vol. 17, pp. 921–924, 2005.

[63] Solymosi, J., "On the Number of Sums and Products," *Bulletin of the London Mathematical Society*, vol. 37, pp. 491–494, 2005.

[64] Solymosi, J., "Bounding Multiplicative Energy by the Sumset," *Advances in Mathematics*, vol. 222, pp. 402–408, 2009.

[65] Solymosi, J., "Incidences and the spectra of graphs," in *Combinatorial Number Theory and Additive Group Theory*, Birkhäuser Basel, 2009.

[66] Solymosi, J. and Tao, T., "An Incidence Theorem in Higher Dimensions," *Discrete and Computational Geometry*, vol. 48, pp. 255–280, 2012.

[67] Székely, L., "Crossing Numbers and Hard Erdős Problems in Discrete Geometry," *Combinatorics, Probability and Computing*, vol. 6, pp. 353–358, 1997.

[68] Szemerédi, E. and Trotter, W. T., "Extremal Problems in Discrete Geometry," *Combinatorica*, vol. 3, pp. 381–392, 1983.

[69] Tao, T., "Expanding Polynomials over Finite Fields of Large Characteristic, and a Regularity Lemma for Definable Sets." `http://terrytao.wordpress.com/2012/11/14/expanding-polynomials-over-finite-fields-of-large-characteristic-and-a-regularity-lemma-for-definable-sets`. Accessed: 2014-09-22.

[70] Tao, T., "The Sum-Product Phenomenon in Arbitrary Rings," *Contributions to Discrete Mathematics*, vol. 4, pp. 59–82, 2009.

[71] Tao, T., "Expanding Polynomials over Finite Fields of Large Characteristic, and a Regularity Lemma for Definable Sets." arXiv:1211.2894, 2013.

[72] Tao, T. and Vu, V. H., *Additive Combinatorics*. Cambridge University Press, 2010.

[73] Vu, V. H., "Sum-Product Estimates via Directed Expanders," *Mathematical Research Letters*, vol. 15, pp. 375–388, 2008.